# Predicting the Pose of β-Casomorphin-5 and 7 in the Opioid Receptors

Christopher Oberc, Bachelor of Science

---

Submitted in partial fulfillment of the requirements for the degree of

---

MASTER OF SCIENCE

---

Faculty of Mathematics & Sciences, Brock University
St. Catharines, Ontario

## Abstract

The opioid receptors consist of three main subtypes; μ, δ, and κ. Previous binding studies have shown that fragments of the milk protein, β-casein, known as β-casomorphins are agonists of these receptors which are selective for the μ receptor subtype. Using the crystal structures of these three receptors, computational molecular docking studies were done using the software GOLD to determine the conformation of β-casomorphin-5 and 7 when they bind to these three opioid receptors. GOLD was able to discriminate among the three receptors when docking the rigid ligands co-crystalized with the receptors. However, GOLD could not discriminate among the three receptors for either of the highly flexible β-casomorphins. A per amino acid scoring method was developed to overcome this problem. This method was used to predict the conformation of both β-casomorphin-5 and 7 in the μ receptor and determine that the two amino acid residues, Lys303 and Trp318 of the μ receptor are responsible for discriminating among the three receptor subtypes for binding of the β-casomorphin-5 and 7.

## Acknowledgements

I would like to thank Heather Gordon and Hongbin Yan for supporting me throughout my graduate degree. I would also like to thank Jeffery Atkinson and Randall Dumont for their contributions. I am also grateful for Michelle Eisner whom has assisted me throughout my degree. Finally, I would like to thank any other faculty members and students that have contributed towards my degree.

## Table of Contents

Abstract ..... i

Acknowledgements ..... ii

Table of Contents ..... iii

List of Figures ..... vii

List of Tables ..... xv

List of Abbreviations ..... xvii

Introduction ..... 1

1: Opioid Receptors ..... 1

1a: Function and Biological Relevance ..... 1

1b: Physiological Function of Opioid Receptors ..... 8

1c: Amino Acid Sequence of the WT Receptors ..... 11

1d: X-Ray Crystal Structures of Proteins ..... 14

1e: Protein Data Bank and Limitations of X-Ray Structures ..... 15

1f: Opioid Receptor-Lysozyme Chimera ..... 16

1g: Opioid Receptor Ligands ..... 24

2: Opioid Peptides ..... 28

2a: Endogenous Opioid Receptor Peptide Ligands ..... 28

2b: Exogenous Opioid Receptor Peptide Ligands ..... 28
3: Computational Molecular Docking ..... 32
3a: Genetic Algorithm ..... 32
3b: Genetic Optimization for Ligand Docking (GOLD) ..... 37
3c: Scoring Function ..... 39
3d: Root Mean Squared Deviation ..... 48
3e: Rescoring ..... 49
Materials and Methods ..... 50
1: Crystal Structure Modifications ..... 51
2: Binding Site Parameters ..... 53
3: GA Parameters ..... 54
4: Flexible Side chains ..... 54
Results ..... 56
1: Receptor Superpositioning ..... 56
2: Native Docking ..... 58
2a: δ Receptor ..... 58
2b: μ Receptor ..... 63
2c: κ Receptor ..... 66
2d: 50Å Radius Docking ..... 68
3: Non-Native Cross Docking ..... 72

3a: Naltrindole ..... 72

3b: β-FNA ..... 73

3c: JDTic ..... 75

4: β-Casomorphin Docking ..... 76

4a: β-Casomorphin-7 ..... 76

4b: β-Casomorphin-5 ..... 80

5: DAMGO Docking ..... 85

5a: Traditional Scoring Analysis ..... 85

5b: Per Amino Acid Scoring Analysis ..... 87

5c: Conclusion Concerning DAMGO Docking ..... 94

6: Per Amino Acid Scoring Analysis of β-Casomorphin-5 ..... 94

6a: Rescoring of Poses of β-casomorphin-5 from the μ Receptor Docking Experiments in δ and κ Receptors ..... 95

6b: Mutant μ Receptor Rescoring ..... 99

7: Per Amino Acid Scoring Analysis of β-Casomorphin-7 Docked to μ Receptor ..... 102

7a: Rescoring μ Poses to δ and κ Receptors ..... 103

7b: μ Mutant Rescoring Analysis ..... 115

Discussion ..... 135

1: Non-Native β-FNA Docking ..... 135

2: DAMGO Pose Analysis ..... 142

3: β-Casomorphin-5 Pose 268 in the W318Y Mutant ..... 143

4: β-Casomorphin-7 Pose Selection ..... 144

5: Global Receptor Conformation ..... 147

Conclusion ..... 151

References ..... 153

Appendix ..... 159

## List of Figures

Figure 1: Chemical structure of morphine and dihydromorphine-[7,8-$^{3}$H]. ..... 2

Figure 2: Structure of MCOPPB ..... 4

Figure 3: Structural determination of rhodopsin. ..... 5

Figure 4: Possible arrangements of the seven TMs viewed from the intracellular surface of the receptor[22]. ..... 8

Figure 5: Schematic representation of the GPCR signal transduction[26]. ..... 10

Figure 6: Chemical structure of DAMGO. ..... 12

Figure 7: Clustal Omega (1.2.0)[33,34] sequence alignment of the wild type (WT) human κ, mouse δ and mouse μ opioid receptors. The TMs are shown in bold. Asterisks indicate conservation of an amino acid for all three receptors at the given position, colons indicate a strong conservation of amino acid properties at the given position and a period indicates a weak conservation of amino acid properties. ..... 13

Figure 8: Amino acid sequences of the κ, δ and μ receptor-lysozyme chimeras. ..... 18

Figure 9: Structure of monoolein. ..... 19

Figure 10: Clustal Omega (1.2.0)[33,34] sequence alignment of the mouse and human WT δ receptors.TMs are shown in bold. ..... 20

Figure 11: Clustal Omega (1.2.0)[33,34] sequence alignment of the mouse and human WT μ receptors.TMs are shown in bold. ..... 21

Figure 12: Structures of the co-crystalized ligands of the opioid receptors. ..... 22

Figure 13: Crystal structures of a) μ opioid receptor b) δ opioid receptor c) κ opioid receptors. 23

Figure 14: Structures of the exogenous ligands from Table 2. ..... 25

Figure 15: Schematic representation of the “chromosome” of butane and its “genes”. ..... 34

Figure 16: Schematic representation of crossovers and mutations. "$R_1$" is the gene that encodes for the dihedral angle of the central C-C bond (Φ) of the butane molecule. Gene "O" encodes for the orientation of the butane (α,β,γ) and gene "P" encodes for the centre-of-mass of butane in space (x,y,z). a) A crossover where the dihedral angle gene ("$R_1$") is exchanged between two chromosomes, b) a crossover where the centre-of-mass gene ("P") is exchanged, and c) a mutation of the orientation gene ("O") causing a random re-orientation of the molecule in 3-dimensional space. ..... 35

Figure 17: Relationship between score of an interaction and distance between interacting atoms of (a) attractive and (b) repulsive atom pairs. ..... 41

Figure 18: H-bonding (dash lines) classification used by CHEMPLP ..... 44

Figure 19: Definition of $p_r$, the $p_\alpha$ and $q_\beta$ angles and the constants for H-bonding functions in CHEMPLP. ..... 45

Figure 20: Metal coordination (dash lines) classification used by CHEMPLP ..... 46

Figure 21: Schematic of the clash cut-off ..... 47

Figure 22: Schematic comparison between docking and rescoring. ..... 49

Figure 23: Superposition of the μ (red), δ (blue) and κ (orchre) receptors: (a) and (b) Viewed from the extracellular face, (c) and (d) viewed from the intracellular face. ..... 57

Figure 24: Comparison of the CHEMPLP score and the RMSD of 100 dockings of neutral naltrindole to the δ receptor using a 30 Å search radius. ..... 59

Figure 25: Comparison of the CHEMPLP score and the RMSD of 100 dockings of protonated naltrindole to the δ receptor using a 30Å search radius. ..... 60

Figure 26: Frequency of occurrence of 100 naltrindole docked poses within various ranges of RMSD to the crystal structure. ..... 61

Figure 27: Highest scoring docked poses (red) of a) neutral naltrindole (score=71.7, RMSD=1.09Å) and b) protonated naltrindole (score=77.9, RMSD=1.10Å) shown with the crystal structure pose (blue). The extra substituents of the docked naltrindole are hydrogen atoms. ..... 62

Figure 28: Salt bridge formation between ammonium centre (proton in cyan) of naltrindole (red) and Asp128. ..... 63

Figure 29: Comparison of the CHEMPLP score and the RMSD of 100 dockings of protonated β-FNA with the μ receptor using a 30 Å search radius. ..... 64

Figure 30: Lowest and highest scoring docked poses of β-FNA ..... 65

Figure 31: Comparison of top docking poses of a) β-FNA (red) docking to the μ receptor with covalent linkage to Lys233 (green) and b) naltrindole (red) docking to the δ receptor with Lys214 (green) (ammonium centre proton in cyan). ..... 66

Figure 32: Comparison of the CHEMPLP score and the RMSD of 100 dockings of protonated JDTic with the κ receptor using a 30 Å search radius. ..... 67

Figure 33: Highest scoring docked poses (red) of JDTic ..... 68

Figure 34: Comparison of docking results using 30 Å and 50 Å search radii. ..... 69

Figure 35: JDTic (green) docked pose that lies outside the main binding pocket of the κ receptor. ..... 71

Figure 36: Comparison of the CHEMPLP score distribution of 100 dockings of naltrindole to the δ, κ and μ receptors. ..... 73

Figure 37: Comparison of the CHEMPLP score distribution of 100 dockings of β-FNA to the δ, κ and μ receptors. ..... 74

Figure 38: Highest scoring pose of β-FNA (red) in a) μ receptor and b) δ receptor (ammonium centre proton in cyan). ..... 74
Figure 39: Comparison of the CHEMPLP score distribution of 100 dockings of JDTic to the δ, κ and μ receptors. ..... 75
Figure 40: 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with rigid receptor side chains. ..... 77
Figure 41: Trial 1, 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with flexible side chains. ..... 78
Figure 42: Trial 1, 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with flexible side chains using cumulative frequency. ..... 79
Figure 43: 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with trial 2 flexible side chains. ..... 80
Figure 44: 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors. ..... 81
Figure 45: 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors with flexible side chains. ..... 82
Figure 46 : Structure of a) L-proline and b) 2-aminocyclopentane-carboxylic acid. ..... 82
Figure 47: Subset of the 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors with flexible side chains bearing a *cis* Tyr-Pro peptide bond. ..... 83
Figure 48: Subset of the 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors with flexible side chains bearing a *trans* Tyr-Pro peptide bond. ..... 84
Figure 49: 500 dockings of DAMGO of the rigid δ, κ and μ receptors. ..... 86
Figure 50: 500 dockings of DAMGO of the δ, κ and μ receptors with flexible side chains. ..... 87

Figure 51: Visual representation of poses a) 463 (CHEMPLP score = 94.67), b) 94 (CHEMPLP score = 85.56), and c) 75 (CHEMPLP score = 90.87) of DAMGO (green) in the μ receptor. ..... 88

Figure 52: Individual amino acid contributions to the scores of the DAMGO μ receptor poses 463, 94 and 75. ..... 89

Figure 53: Per residue scores from the rescoring of DAMGO poses a) 463, b) 94, and c) 75 in the μ, δ and κ receptors. ..... 91

Figure 54: Visual representations of μ receptor DAMGO (green) poses 463 and 94 rescored in the δ and κ receptors. ..... 93

Figure 55: Individual amino acid contributions of the scores of the β-casomorphin-5 μ receptor poses 46, 2 and 468. ..... 95

Figure 56: Residue scores from rescoring of the β-casomorphins μ receptor poses a) 46, b) 2, and c) 268 in the μ, δ and κ receptors. ..... 96

Figure 57: Visual representations of the rescoring of β-casomorphin-5 pose 268 in the a) μ receptor, b) δ receptor, and c) κ receptor. ..... 98

Figure 58: Per residue scores of the β-casomorphin-5 μ receptor pose 268 rescored in K303 and W318 μ receptor mutants. ..... 99

Figure 59: Visual representations of β-casomorphin-5 μ receptor pose 268 rescored in μ receptor. ..... 101

Figure 60: Individual amino acid contributions towards the scores of the μ receptor poses 468, 9, 235 and 12. ..... 103

Figure 61: Residue scores from rescoring of the μ receptor poses a) 468, b) 9, c) 235, and d) 12 in the δ and κ receptors. ..... 105

Figure 62: Visual representation of β-casomorphin-7 pose 468 in the μ receptor rescored in a-b) μ receptor, c-d) δ receptor, and e-f) κ receptor. ..... 107
Figure 63: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a-b) μ receptor, c-d) δ receptor, and e-d) κ receptor. ..... 110
Figure 64: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a-b) μ receptor, c-d) δ receptor, and e-d) κ receptor. ..... 112
Figure 65: Visual representation of β-casomorphin-7 (green) pose 12 in the μ receptor rescored in a) μ receptor, b) δ receptor, and c) κ receptor. ..... 115
Figure 66: Residue scores from rescoring of the β-casomorphin-7 μ receptor pose 468 in mutant μ receptors. ..... 116
Figure 67: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) N127K μ receptor, and c) N127V μ receptor. ..... 117
Figure 68: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor and b) V300I μ receptor. ..... 118
Figure 69: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) K303W μ receptor, and c) K303E μ receptor. ..... 119
Figure 70: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) E310R μ receptor, and c) E310H μ receptor. ..... 121
Figure 71: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor. ..... 122
Figure 72: Residue scores from rescoring of the μ receptor pose 9 in mutant μ receptors. ..... 123
Figure 73: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor, b) N127K μ receptor, and c) N127V μ receptor. ..... 124

Figure 74: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor and b) N230L μ receptor. ..... 125
Figure 75: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor and b) V300I μ receptor. ..... 125
Figure 76: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor. ..... 126
Figure 77: Residue scores from rescoring of the μ receptor pose 235 in mutant μ receptors. ... 127
Figure 78: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor, b) N127K μ receptor, and c) N127V μ receptor. ..... 128
Figure 79: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor and b) E229D μ receptor. ..... 129
Figure 80: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor and b) V300I μ receptor. ..... 130
Figure 81: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor. ..... 131
Figure 82: Residue scores from rescoring of the μ receptor pose 12 in mutant μ receptors. ..... 132
Figure 83: Visual representation of β-casomorphin-7 (green) pose 12 in the μ receptor rescored in a) WT μ receptor and b) N230L μ receptor. ..... 133
Figure 84: Visual representation of β-casomorphin-7 (green) pose 12 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor. ..... 134
Figure 85: β-FNA a) Rescoring of top scoring μ receptor pose in the κ receptor, b) Top scoring β-FNA pose in δ receptor, and c) Top scoring β-FNA pose in μ receptor. ..... 136
Figure 86 Comparison of unreacted β-FNA in the μ verses δ receptor. ..... 139

Figure 87: Quantitative analysis of top scoring β-FNA pose in the a) μ and b) δ receptors. ..... 141

Figure 88: Superposition of W318Y μ receptor (red) and WT κ receptor (blue). ..... 144

Figure 89: Superposition of pose 268 of β-casomorphin-5 (green scaffold) with a) pose 468, b) pose 12, c) pose 235, and d) pose 9 of β-casomorphin-7 (black scaffolds). ..... 145

Figure 90: Structures of alkaloid ligands used. ..... 148

Figure 91: Distribution of the three poses of morphine in the μ receptor. ..... 149

Figure 92: Comparison of the CHEMPLP score distribution of 100 dockings of morphine in the μ, δ and κ receptors. ..... 150

Figure 93: Superpositioning of naltrindole and morphine a) Comparison of the docked pose of morphine (black scaffold) in the δ receptor and the crystal structure pose of naltrindole (green scaffold), RMSD=0.58Å and b) Atoms used for an RMSD calculation (black). 150

## List of Tables

Table 1: Dissociation constants (nM) of various compounds that were displaced by naloxone and Mr 2266[8]. ..... 3

Table 2: Exogenous ligands and their binding affinities for the δ, μ and κ receptors[49]..... 24

Table 3: Endogenous ligands of the opioid receptors. ..... 28

Table 4: $IC_{50}$ (μM) values of exogenous opioid receptor peptide ligand towards guinea pig ileum (GPI) which contains mainly μ receptors and mouse vas deferens (MVD) which contains mainly δ receptors[57]. ..... 29

Table 5: $IC_{50}$ values (μM) to the GPI and MVD. ..... 30

Table 6: A list of the ligands with their $IC_{50}$ values (μM) to homogenized rat brain. ..... 31

Table 7: Undetected or missing whole amino acids and amino acid side chains in opioid receptors PDB files. ..... 51

Table 8: Tautomeric forms of histidines used and rationale for choosing them. ..... 53

Table 9: Identities of the central atoms of the binding site search radii used for all docking experiments. ..... 54

Table 10: List of the GA parameters. ..... 54

Table 11: List of amino acid side chains in receptor binding site allowed to be flexible. ..... 55

Table 12: RMSD values of the backbone atoms (N, $C_\alpha$, C and O) of each of the receptor pairs. The terminal regions of the sequences were not included in the calculation because of their varying amino acid lengths. ..... 56

Table 13: Ratio of *cis* and *trans* Tyr-Pro peptide bond among the δ, κ and μ receptors. ..... 83

Table 14: List of the amino acids of the four poses, 468, 9, 235 and 12 predicted to the responsible for discriminating β-casomorphin-7 between the three receptors. ..... 135

Table 15: The comparison of the RMSD of various components of β-casomorphin-5 and 7. Row 3 compares the backbone atom (N, $C_\alpha$, C and O) of the Phe-Pro-Gly sequence of β-casomorphin-5 and the Gly-Pro-Ile sequence of β-casomorphin-7. ..... 146

## List of Abbreviations

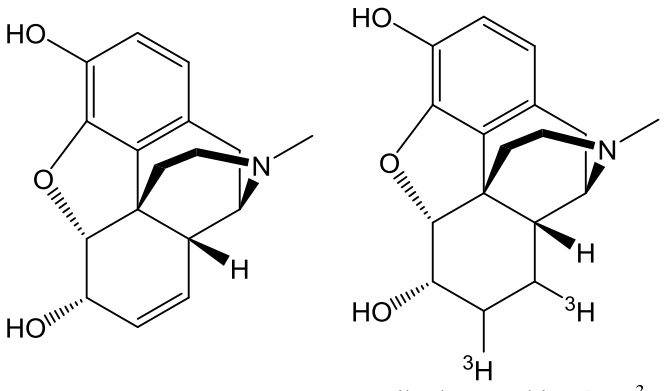
| | |
|---|---|
| ACPC | 2-Aminocyclopentane Carboxylic Acids |
| ATP | Adenosine 5'-Triphosphate |
| cAMP | Adenosine 5',3'-Monophosphate |
| CNA | Chlornaltrexamine |
| DADLE | Tyr-D-Ala-Gly-Phe-D-Leu |
| DAMGO | Tyr-D-Ala-Gly-N-Me-Phe-Gly-ol |
| ECL | Extracellular Loop |
| GA | Genetic Algorithm |
| GDP | Guanosine 5'-Diphosphate |
| GOLD | Genetic Optimization for Ligand Docking |
| GPCR | G-Protein Coupled Receptor |
| GPI | Guinea-Pig Ileum |
| GppNHp | Guanosine 5′-[β,γ-imido]triphosphate |
| GTP | Guanosine 5'-Triphosphate |
| GTPase | Guanosine 5'-Triphosphatase |
| H-bonds | Hydrogen Bonds |
| ICL | Intracellular Loop |
| JDTic | (3R)-7-Hydroxy-N-[(2S)-1-[(3R,4R)-4-(3-Hydroxyphenyl)-3,4-Dimethylpiperidin-1-yl]-3-Methylbutan-2-yl]-1,2,3,4-Tetrahydroisoquinoline-3-Carboxamide |
| MD | Molecular Dynamics |
| MVD | Mouse Vas Deferens |
| N/OFQ | Nociceptin/Orphanin FQ |
| OR | Opioid Receptor |
| PDB | Protein Data Bank |
| PLP | Piecewise Linear Potential |
| RMSD | Root Mean Square Deviation |
| RVD | Rat Vas Deferens |
| T4L | T4 Lysozyme |
| TEV | Tobacco Etch Virus |
| TM | Transmembrane Helix |
| WT | Wild Type |
| β-AR | $\beta_2$-Adrenergic Receptor |
| β-FNA | β-Funaltrexamine |

# Introduction

## 1: Opioid Receptors

### 1a: Function and Biological Relevance

Morphine and its analogues are used medically as potent analgesics[1]. Morphine was first isolated from opium in 1804 by Friedrich Wilhelm Sertürner[2]. Despite this early achievement, it was not until 1966 that the receptors to which morphine and its analogues bind were identified. Praag and Simon identified the brain as having the highest concentration of the receptors by adding dihydromorphine-[7,8-$^{3}$H] (a compound that induces similar effects to morphine) (Figure 1) to homogenized rat organs[3]. Later, Terenius homogenized guinea-pig ileum (GPI) and treated the homogenate with dihydromorphine-[7,8-$^{3}$H][4,5]. The different components of the cells (nuclei, mitochondria and plasma membrane) were fractionated. Most of the radioactivity was found in the plasma membrane fraction, showing that the dihydromorphine receptors were membrane-bound. These experiments did not indicate how many subtypes of the receptor there were, nor did they show whether the receptor is simply tethered to the membrane or fully embedded in it. By this time, the receptor had been named the opioid receptor (OR) for its affinity for opioid alkaloids[6].




Morphine Dihydromorphine-[7,8-$^3$H]

**Figure 1: Chemical structure of morphine and dihydromorphine-[7,8-$^3$H].**

The μ and κ receptor subtypes were identified by Martin and co-workers in 1976[7]. Their experiments involved making dogs dependent on morphine, a known opioid receptor agonist. The results indicated that the morphine-dependent dogs still experienced withdrawal symptoms when morphine was substituted by ketocyclazocine (an alkaloid with a similar structure to morphine). Likewise, when dogs were made dependent on ketocyclazocine they experienced withdrawal symptoms when given morphine instead. It was noted that the symptoms of morphine and ketocyclazocine withdrawal were different. The only explanation for this observation was that these two opioids act on different receptors[7].

The δ receptor was discovered in 1977 by Lord and co-workers[8]. They performed displacement binding experiments on mouse vas deferens (MVD) (Table 1).

**Table 1: Dissociation constants (nM) of various compounds that were displaced by naloxone and Mr 2266**[8].

| Compound | Naloxone | Mr 2266 |
|---|---|---|
| Normorphine | 1.84±0.20 | 1.50±0.14 |
| Met-enkephalin | 22.6±0.64 | 16.2±1.5 |
| Leu-enkephalin | 21.4±3.3 | 34.0±5.3 |
| Ethylketocyclazocine | 11.0±0.6 | 4.51±0.54 |
| Mr 2034 | 9.1±0.7 | 8.4±1.3 |

It was already known that morphine and naloxone were μ-selective ligands and that ethylketocyclazocine (a derivative of ketocyclazocine) was a κ-selective ligand based on the work of Martin[7]. The penta-peptides Met-enkephalin and Leu-enkephalin were known to be two of the endogenous ligands of the receptors[9]. Mr 2266 is a κ-selective antagonist and Mr 2034 is a κ-selective agonist[10,11]. The MVD could undergo contractions when stimulated by an electric pulse and the extent of this contraction is inhibited by agonists of opioid receptors but not antagonists[12]. Therefore, since all the compounds in Table 1: column 1 are agonists, and both compounds in Table 1: row 1 are antagonists, the binding affinity can be measured by how much antagonist is required in order for the MVD to have the same contraction as the control. Higher concentrations of both naloxone and Mr 2266 were required to displace these peptide ligands (see Table 1: rows 2 and 3) than the known μ- and κ-selective ligands morphine and ethylketocyclazocine, respectively, indicating that there was a third receptor subtype[8].

The fourth opioid receptor, nociceptin/orphinan (N/OFQ) was discovered by Bunzow and co-workers[13]. They performed a sequence homology study on fragments of the mouse genome using the mouse δ opioid receptor gene as a reference. A gene bearing considerable amino acid sequence similarity (47% overall) was identified. Further analysis showed the same overall

amino acid sequence similarity between the N/OFQ receptor and the mouse μ and κ opioid receptors. The transmembrane portions of the N/OFQ receptor have a higher amino acid sequence similarity (64%) to the δ receptor. Binding assays showed that this receptor does not bind well to ligands known to be selective for one of the other receptors (μ, δ or κ)[13]. The endogenous ligand of this receptor is nociceptin, a heptadeca-peptide[14]. MCOPPB (Figure 2) is one of the exogenous ligands that is selective for this receptor[15].

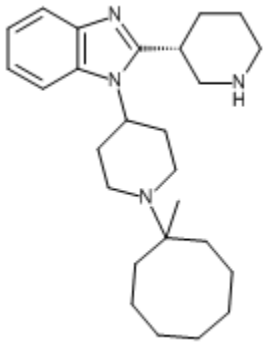



**Figure 2: Structure of MCOPPB**

In 1980, Hargrave and co-workers determined the structural features of rhodopsin. Bovine cells bearing rhodopsin were treated externally with a variety of peptidases[16]. After the peptidase exposure, the cell membrane was lysed with a detergent and the fragments of the receptor were analyzed by gel filtration chromatography using a size exclusion resin. Amino acid segments of the receptor embedded in the membrane were inaccessible to the peptidases, thus never experienced peptide bond cleavage. When a carboxypeptidase (a peptidase that degrades peptides at the C-terminus) was used, no degradation was observed. The conclusion from this was that the C-terminus was on the intracellular face of the cell. Freezing and thawing the membrane can lead to a flip of the orientation of the receptor[17]. When thawed membranes were exposed to the carboxypeptidase, some degradation was seen. This verified that the C-terminus

was intracellular in the native orientation. The degradation patterns of the receptor in the normal and flipped orientations were analysed after a variety of peptidases that cleave internal peptide bonds were used. It was concluded that the receptor contained seven transmembrane helices (TMs) based on the observation of seven peaks observed from the chromatography. However, the researchers were not completely certain of the presence of seven helices as it was possible for some peaks to come from large fragments of the ECLs, ICLs or termini. A helical conformation of the transmembrane portions of the sequence (*ca.* 40 amino acids each, 5.4 nm length) was posited because this was consistent with a length equal to that of the width of the plasma membrane (4 nm). Given that the C-terminus is intracellular and that the receptor has seven TMs (Figure 3), the N-terminus must be extracellular[16]. This conclusion was reached because in order for the seven helices of the single peptide chain to be connected, three extracellular loops (ECL) and three intracellular loops (ICL) are required (Figure 3). The consequence of this is that the N-terminus will be on the extracellular surface (Figure 3).

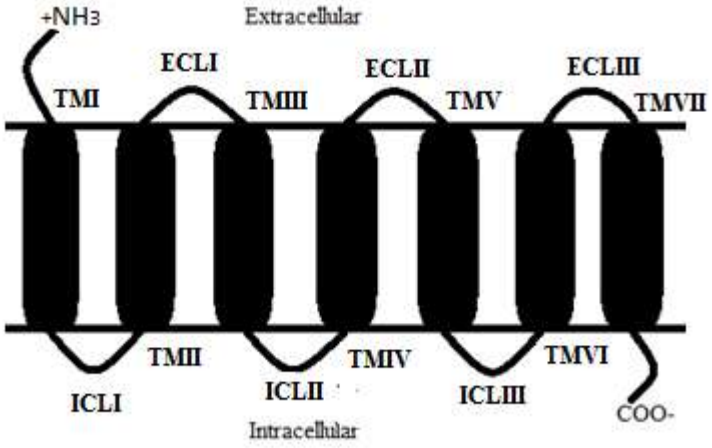



**Figure 3: Structural determination of rhodopsin. The black cylinders represent the TMs and the two horizontal lines represent the plasma membrane[16].**

The extent of amino acid sequence similarity between rhodopsin and the opioid receptors (μ: 24%, δ: 22% and κ: 20%) led researchers to hypothesize that opioid receptors have the same structural features, especially considering that most of the sequence similarity is localized in the proposed helical regions[18].

In 1986, Graham and co-workers determined that the $\alpha_2$-adrenergic receptor is coupled to a G-protein. They used extracted rat membrane and measured the activity of the known antagonist prazosin ($^3$H-labeled) over the agonist epinephrine (adrenaline) (unlabeled) and introduced the unhydrolyzable analogue of guanosine-5'-triphosphate (GTP), guanosine-5′-[β,γ-imido]triphosphate (GppNHp). The measured $IC_{50}$ of the prazosin was dependent on whether or not a protease inhibitor and GppNHp were added. The presence of the protease inhibitor resulted in a decrease of prazosin binding while no difference in prazosin binding was seen in the absence of the GppNHp whether or not a protease inhibitor was present. The conclusion of this experiment was that the $\alpha_2$-adrenergic receptor binds to a G-protein (which then binds to a GTP) in the presence of an agonist bound to the receptor but not an antagonist[19].

In 1988, Costa and co-workers showed that addition of the selective opioid receptor alkylating agent β-chlornaltrexamine (CNA) to a cell culture will result in a reduced guanosine-5'-triphosphatase (GTPase) activity relative to the control culture treated with naltrexone[20]. Higher concentrations of CNA resulted in greater reductions in GTPase activity. Treating the cell culture with pertussis toxin, a G-protein antagonist, had similar results. These observations suggested that the opioid receptors were also GPCRs. The following year, Allgaier and co-workers performed a series of experiments on rabbit hippocampal tissue which contained both $\alpha_2$-adrenergic receptors and κ opioid receptors[21]. They found that the $IC_{50}$ of the competitive binding of $^3$H-noradrenaline and yohimbine, an agonist and antagonist, respectively, of the

adrenergic receptor is dependent on the activity of the κ opioid receptors. When ethylketocyclazocine was added in addition to the $^{3}$H-noradrenaline and yohimbine, the $IC_{50}$ decreased (less noradrenaline and more yohimbine bound to the adrenergic receptor). The conclusion was that when each of these receptors had an agonist bound, they competed for the same signalling molecule, which was too low in concentration *in vivo* to saturate both receptors simultaneously. This molecule was believed to be the G-protein, as the work done by Graham showed that the G-protein interacted with the adrenergic receptors[19]. When pertussis toxin was used, the presence or absence of ethylketocyclazocine had no effect on the $IC_{50}$ between noradrenaline and yohimbine, further supporting an interaction between the G-proteins and both of the receptors[21].

The architecture of rhodopsin deduced by Hargrave[16] was further elaborated on by Baldwin in 1993[22]. Baldwin took the sequences of many different GPCRs. The seven transmembrane segments were identified based on the high percentage of hydrophobic amino acids. It was observed that all the helices had segments that were completely hydrophobic and segments that were amphipathic. It was proposed that the hydrophobic segments face the membrane and that the amphipathic segments interact with other helices. This analysis showed that TMI has a larger hydrophobic surface area than those of the other six helices, therefore it is likely the most exposed to the membrane, while TMIII is more polar and thus should have limited contact with the membrane and more contact with other helices. The remaining five TMs have intermediate levels of membrane exposure based on this analysis. This resulted in two possible arrangements of the helices (Figure 4). The only difference between the two arrangements is the directionality of the helices (clockwise verses counter-clockwise) which

makes these two arrangements mirror images of each other. This arrangement held true for all the GPCR sequences tested[22].

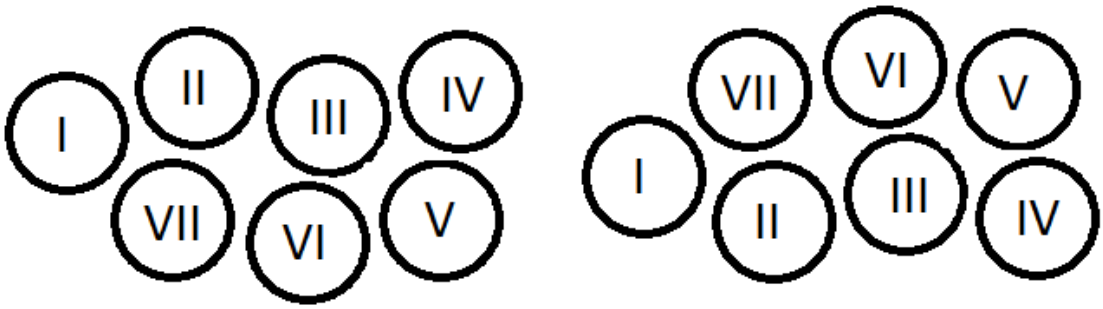



**Figure 4: Possible arrangements of the seven TMs viewed from the intracellular surface of the receptor[22].**

In 2000, Palczewski and co-workers resolved the crystal structure of bovine rhodopsin[23]. This crystal structure has been used by other researchers for modelling opioid receptors, however, these models were very inaccurate with respect to opioid receptor ligand binding since the amino acid identities between bovine rhodopsin and the opioid receptors are less than 30%[18].

**1b: Physiological Function of Opioid Receptors**

The four opioid receptors, μ, δ, κ and N/OFQ have different physiological roles. The δ receptor is named after the (vas) deferens, the first organ from which the receptor was isolated. It causes analgesia, which can influence gastrointestinal motility, respiration, cognitive functions and mood[24,25]. The κ receptor acquired its name from ketocyclazocine, a ligand known to be highly selective for binding to the κ receptor. This receptor affects feeding which can further lead to changes in diuresis and neuroendocrine secretions[24,25]. The μ receptor is named after morphine, which has the highest binding affinity for the μ over the δ and κ receptors. This receptor causes analgesia which can further lead to changes in respiration, cardiovascular functions and immune functions[24,25]. The N/OFQ receptor was named after being identified as an

orphan opioid receptor via homology cloning[13]. Furthermore, the endogenous ligand nociceptin binds to the N/OFQ receptor with high binding affinity[14]. It causes hyperalgesia which can influence the immune functions, gastrointestinal tract and heart rate. It is also believed to play a role in addiction to exogenous ligands that bind primarily to other opioid receptors (δ, μ and κ)[14].

The signal transduction of GPCR starts with an agonist ligand binding to the receptor on the extracellular interface[26]. This results in a conformational change that recruits the G-protein. The G-protein in its inactive form is a heterotrimer composed of an α, β and γ subunit with a guanosine-5'-diphosphate (GDP) bound to the α subunit. Upon the recruitment of this heterotrimer to the GPCR, the α subunit binds to the intracellular interface of the GPCR and dissociates from the β and γ subunits. These two subunits remain as a heterodimer. When the α subunit is bound to the GPCR, the GDP bound to it will be displaced by a GTP; dissociation of the subunit from the receptor follows. This Gα-GTP complex can trigger various cascades. Gα-GTPase will eventually dephosphorylate the bound GTP to create GDP, which results in the reassociation of the βγ heterodimer to the α subunit, reforming the inactive heterotrimer[26] (Figure 5).

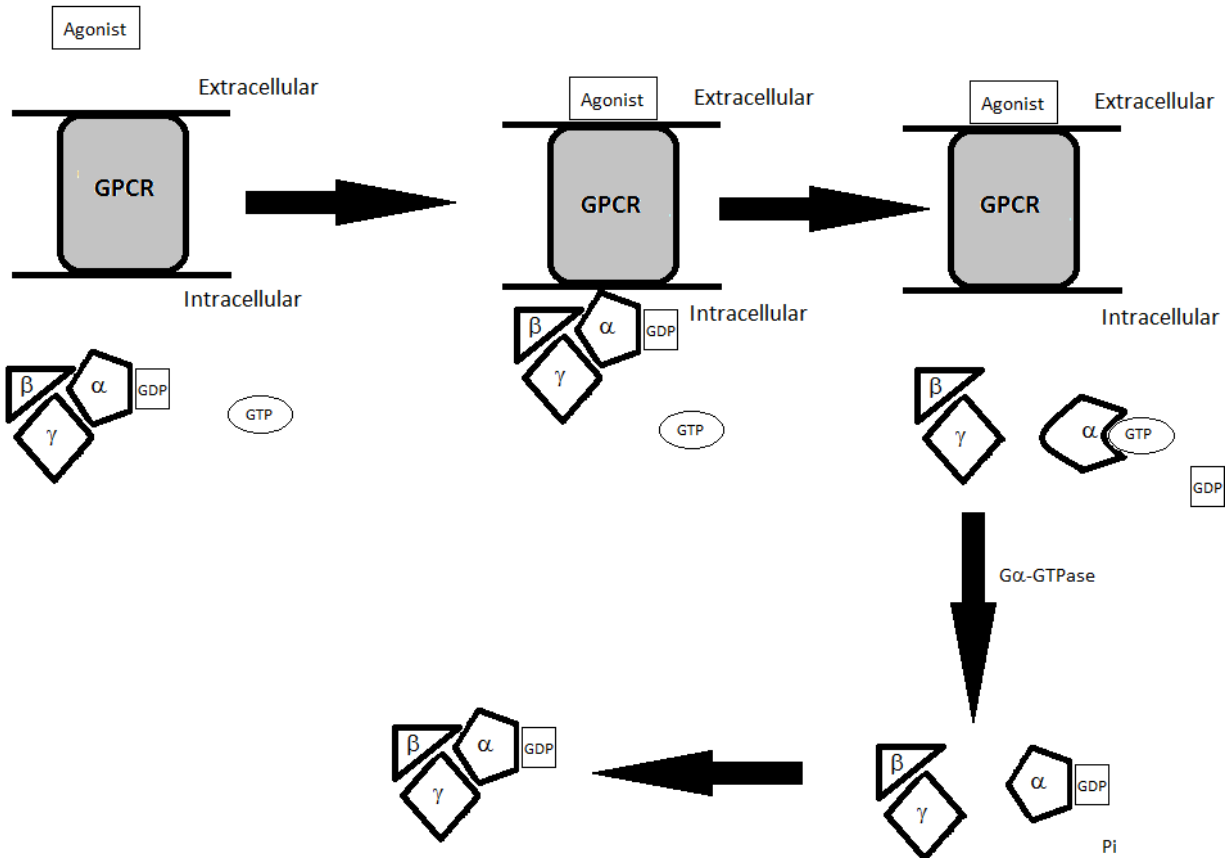



**Figure 5: Schematic representation of the GPCR signal transduction[26].**

One of the cascades that the Gα-GTP complex can influence is the synthesis of cyclic adenosine 3',5'-monophosphate (cAMP)[27]. In this cascade, the Gα-GTP complex binds to adenylate cyclase which is also a membrane-bound receptor. However, unlike the opioid receptors, adenylate cyclase contains 12 TMs and an intracellular domain between helices six and seven[28]. By default, adenylate cyclase is active in converting adenosine 5'-triphosphate (ATP) into cAMP and pyrophosphate. The cAMP is a signaling molecule for other cascades. Upon binding of the Gα-GTP complex, the adenylate cyclase will adopt the inactive conformation and thus not produce cAMP[27].

The opioid receptors are also not the only receptors that can activate the G-protein. Another receptor that can do this is the $\beta_2$-adrenergic receptor plus many others. This receptor also falls under the category of GPCRs and likewise has the same seven-transmembrane helical architecture[29]. The principal endogenous ligand for the $\beta_2$-adrenergic receptor is epinephrine (adrenaline)[27].

**1c: Amino Acid Sequence of the WT Receptors**

The three opioid receptors, examined in this thesis, the δ, κ and μ receptors, have highly conserved amino acid sequences of their seven TMs, with sequence similarity of 73% between the μ and κ receptors, 74% between the δ and κ receptors, and 76% between μ and δ receptors[30]. In addition, their three ICL and the ECLI are also highly conserved in their sequences (Figure 7). However, the ECLII and ECLIII are very diverse in their sequences. Further, the extracellular N-termini and intracellular C-termini vary greatly, not only with respect to their sequences but also in their lengths.

The variations in the ECLII and ECLIII are what give the receptors selectivity for specific ligands. An example of the involvement of the ECLIII in receptor discrimination was shown by Seki and co-workers[31]. The group experimented with the synthetic peptide ligand Tyr-D-Ala-Gly-N-Me-Phe-Gly-ol (DAMGO) (Figure 6), an antagonist that binds selectively to the μ receptor. They determined that making specific mutations in the ECLIII (E297K, S310V, Y312W, Y313H) of the κ receptor can increase its affinity for DAMGO to that of the μ receptor[31]. However, mutations in this loop of the δ receptor did not influence the binding affinity of DAMGO. Rather a single mutation of a lysine found at the junction between the TMII and ECLI to an asparagine (K108N) did achieve similar binding to the μ receptor. Mutating K108 to any other amino acid except glycine and tryptophan also showed good binding affinity

for DAMGO. It was also observed that mutants having substitutions of K108 by uncharged amino acids showed higher binding affinity for DAMGO than when charged amino acids replaced K108[32].

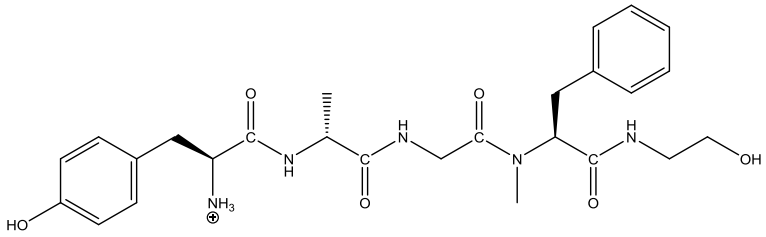



**Figure 6: Chemical structure of DAMGO.**


**Figure 7: Clustal Omega (1.2.0)[33,34] sequence alignment of the wild type (WT) human κ, mouse δ and mouse μ opioid receptors. The TMs are shown in bold. Asterisks indicate conservation of an amino acid for all three receptors at the given position, colons indicate a strong conservation of amino acid properties at the given position and a period indicates a weak conservation of amino acid properties.**

**1d: X-Ray Crystal Structures of Proteins**

Despite the benefits of using morphine as an analgesic, addiction is a major problem associated with its overuse and especially that of its derivative acetylmorphine (heroin). Opiates like these bind primarily to the μ receptor; however, they also bind to the other receptors with some affinity, which is what is believed to result in addiction[35]. Furthermore, this lack of binding specificity leads to side effects such as depression and respiratory arrest. Designing new ligands with higher selectivity for a single receptor subtype is of interest, as these are anticipated to be less addictive than the current opiates on the market and have fewer side effects[36]. Having access to structural information of the receptors would be a major asset into designing such ligands.

The most common method for determining the three-dimensional structure of proteins is X-ray crystallography (X-ray diffraction). This technique requires single crystals of the protein. Single crystal growth involves dissolving the protein in a solution containing various buffers and salts, followed by slow evaporation of the sovent[37]. Different proteins require different conditions in order to form an appropriate crystal for X-ray diffraction. This technique is suited for globular proteins (water soluble) but not the opioid receptors. These proteins, being membrane-bound, have a hydrophobic exterior and will not dissolve in standard crystallization solutions. More importantly, their three-dimensional structures cannot be maintained in an aqueous environment.

The X-ray diffraction method works by bombarding the crystal of the protein with X-ray radiation. When the X-ray photons hit atoms in the protein, they are diffracted and detected by some form of detector[37]. In diffraction experiments, the X-rays are diffracted primarily by the electron clouds around the nuclei[37]. The consequence of this is that hydrogen atoms, which have a low electron density, provide poor diffraction of the X-rays, and thus are not detected[38] by

standard methods. However, high resolution techniques exist that can be employed to detect to locations of hydrogen atoms. These techniques are usually employed when there is a need to determine the position of certain hydrogen atoms[39].

**1e: Protein Data Bank and Limitations of X-Ray Structures**

Computer visualization programs like SwissPDB[40] can display the three-dimensional structure of molecules by reading files that are in the Protein Data Bank (PDB)[41] format. This file format contains information about every atom solved for in the X-ray diffraction experiment. Such information includes three-dimensional atomic coordinates (x, y, z), atom type (carbon, oxygen, nitrogen, etc.), atom identity (N, $C_\alpha$, $C_\beta$, etc.) and amino acid residue (e.g. Trp318). Using this information, SwissPDB will display appropriate bonding between the atoms. SwissPDB also has the ability to introduce hydrogen atoms since their coordinates are normally not detected by X-ray diffraction. In order to predict the coordinates of hydrogen atoms, SwissPDB analyses the unfulfilled valency of each atom and adds the appropriate number of hydrogen atoms. The hydrogen atoms are positioned according to the hybridization state of the atom to which they are bound. For example, SwissPDB will add one hydrogen atom to each $C_\alpha$ (two for glycine) and position them such that tetrahedral geometry is achieved. For the four aromatic hydrogens of tyrosine, SwissPDB will add one hydrogen atom to each $C_\delta$ and $C_\varepsilon$ such that the final geometry is planar. However there are limitations[38]. For example, the protonation state of the imidazole ring of histidine needs to be determined. If it is in its protonated form, both nitrogen atoms of the ring ($N_\delta$ and $N_\varepsilon$) would have a proton; however, if it is in its neutral form then one nitrogen atom of the imidazole ring is protonated and the other nitrogen atom is not. In this scenario, a decision must be made as to which nitrogen is protonated. The decision is usually

made based on optimal hydrogen bond formation of the histidine to nearby amino acids and/or ligands.

Sometimes some of the side chains of amino acids or even entire residues are not detected in the X-ray diffraction experiment. This arises when different conformations of an amino acid in the crystal exist[42]. SwissPDB uses an algorithm to construct locations of missing side chains by positioning the atoms in such a conformation as to minimize the steric clashes with neighbouring amino acid atoms whose coordinates are present in the PDB file. However, SwissPDB cannot construct entire amino acids[40]. For this, other programs, such as ArgusLab[43] can be used. ArgusLab permits the addition of amino acids and calls an algorithm to minimize any steric interactions between the constructed segment(s) and the pre-existing amino acids[43].

**1f: Opioid Receptor-Lysozyme Chimera**

In 2012, the X-ray structures of all four opioid receptors were resolved. Kobilka and co-workers reported the crystal structures of the δ and μ receptors[30,44], while Stevens and co-workers determined the crystal structures of the κ and N/OFQ receptors[45,46]. These structures allow researchers to study the opioid receptors based on structures of the actual receptors rather than models based on related receptors such as rhodopsin.

As mentioned previously (section 1d), standard crystallization techniques that are used for globular proteins are not applicable to the opioid receptors since they are membrane bound proteins. Their hydrophobic exteriors will result in the receptors aggregating *in vitro* instead of dissolving in the aqueous solvent system.

Previous work on forming single crystals of the $\beta_2$-adrenergic receptor (β-AR), another GPCR, for X-ray diffraction showed that crystals could not be formed[29]. The main problem was

in the TMV and TMVI as well as the ICLIII that joins them. This region is considerably more flexible that the other TMs. To counter this problem, T4 lysozyme (T4L) was inserted into the ICLIII to help rigidify both the ICLIII and the neighbouring helices[47]. Lysozyme is a soluble protein known to fold rapidly into its stable folded conformation. The β-AR-T4L chimera was successfully crystallized for X-ray diffraction. The same approach was used to produce structures of the four opioid receptors. The recombinant chimeras of the opioid receptors were expressed by inserting T4L between TMV and TMVI (Figure 8), which were crystallized and subjected to X-ray diffraction.

κ Opioid Chimera

(N-Terminus) (TMI) (ICLI) (TMII)
GGTTMGSEDA QLEPAHISP**A IPVIITAVYS VVFVVGLVGN SLVMFVIIR**Y TKMKT**ATNIY IFNLALADAL VTTTMPFQST** 80
(ECLI) (TMIII) (ICLII) (TMIV)
**VYLM**NSWPFG **DVLCKIVLSI DYYNMFTSIF TLTMMSVDRY IAV**CHPVKAL DF**RTPLKAKI INICIWLLSS SVGISAIV**LG 160
(ECLII) (TMV) (ICLIII/T4L)
GTKVREDVDV IECSLQFPDD D**YSWWDLFMK ICVFIFAFVI PVLIIIVCYT LMILRLKSVR LL**SG*NIFEML RIDEGLRLKI* 240
(T4L)
*YKDTEGYYTI GIGHLLTKSP SLNAAKSELD KAIGRNTNGV ITKDEAEKLF NQDVDAAVRG ILRNAKLKPV YDSLDAVRRA* 320
(T4L) (ICLIII)
*ALINMVFQMG ETGVAGFTNS LRMLQQKRWD EAAVNLAKSR WYNQTPNRAK RVITTFRTGT WDAY*REKDR**N LRRITRLVLV** 400
(TMVI) (ECLIII) (TMVII) (C-Terminus)
**VVAVFVVCWT PIHIFILVEA L**GSTSHSTAA L**SSYYFCIAL GYTNSSLNPI LYAF**LDENFK RCFRDFCFPL KMRMERQSTS 480

δ Opioid Chimera

(N-Terminus) (TMI) (ICLI) (TMII) (ECLI)
GSPGAR**SASS LALAIAITAL YSAVCAVGLL GNVLVMFGIV R**YTKLKT**ATN IYIFNLALAD ALATSTLPFQ SAKYLM**ETWP 80
(TMIII) (ICLII) (TMIV) (ECLII)
FGE**LLCKAVL SIDYYNMFTS IFTLTMMSVD RYIAV**CHPVK ALDFRTP**AKA KLINICIWVL ASGVGVPIMV** MAVTQPRDGA 160
(TMV) (ICLIII/T4L)
VVCMLQFPSP S**WYWDTVTKI CVFLFAFVVP ILIITVCYGL MLLR**LRSVR*N IFEMLRIDEG LRLKIYKNTE GYYTIGIGHL* 240
(T4L)
*LTKSPSLNAA KSELDKAIGR NTNGVITKDE AEKLFNQDVD AAVRGILRNA KLKPVYDSLD AVRRAALINM VFQMGETGVA* 320
(T4L) (ICLIII) (TMVI)
*GFTNSLRMLQ QKRWDEAAVN LAKSRWYNQT PNRAKRVITT RTGTWDAY*EK DRSLR**RITRM VLVVVGAFVV CWAPIHIFVI** 400
(ELCIII) (TMVII) (C-Terminus)
**VWT**LVDINRR D**PLVVAALHL CIALGYANSS LNPVLY**AFLD ENFKRCFRQL CRTPCGRQEP 460

μ Opioid Chimera

(N-Terminus) (TM1) (ICL1) (TMII)
GSHSLCPQTG SPSM**VTAITI MALYSIVCVV GLFGNFLVMY VIVR**YTKMKT **ATNIYIFNLA LADALATSTL PFQSVNYLM**G 80
(ECLI) (TMIII) (ICLII) (TMIV)
TWPFG**NILCK IVISIDYYNM FTSIFTLCTM SVDRYIAV**CH PVKALDFRTP **RNAKIVNVCN WILSSAIGLP VMFM**ATTKYR 160
(ECLII) (TMV) (ICLIII/T4L)
QGSIDCTLTF SHPT**WYWENL LKICVFIFAF IMPVLIITVC YGLMILR**LKS VR*NIFEMLRI DEGLRLKIYK NTEGYYTIGI* 240
(T4L)
*GHLLTKSPSL NAAKSELDKA IGRNTNGVIT KDEAEKLFNQ DVDAAVRGIL RNAKLKPVYD SLDAVRRAAL INMVFQMGET* 320
(T4L) (ICLIII) (TMVI)
*GVAGFTNSLR MLQQKRWDEA AVNLAKSRWY NQTPNRAKRV ITTFRTGTWD AY*EKDRNLR**R ITRMVLVVVA VFIVCWTPIH** 400
(ECLIII) (TMVII) (C-Terminus)
**IYVIIKA**LIT IPET**TFQTVS WHFCIALGYT NSCLNPVLYA F**LDENFKRCF REFCIPTSST IEQP 464

**Figure 8: Amino acid sequences of the κ, δ and μ receptor-lysozyme chimeras. The lysozyme component is shown in italics and TMs are shown in bold.**

In addition to the insertion of the T4L, several other modifications were made to the amino acid sequence[30,44-46]. First, the N and C-terminal regions were truncated. Second, a FLAG sequence was added to the N-terminus and linked to the receptor by a tobacco etch virus (TEV)

protease recognition sequence. Third, a polyhistidine tag was added to the C-terminus. The rationale for truncating the termini was that the terminal regions do not fold into a well-defined structure and as a result, interfere with the crystallization of the chimera.

The opioid receptor-T4L chimeras were expressed in host Sf9 cells, transfected by a baculovirus. The cells were lysed and the lysate was passed through nickel-nitrilotriacetic acid agarose. This resin binds strongly to polyhistidine tags, thus retaining the chimera on the resin. Proteins lacking this tag are washed away. The bound chimeric proteins were then released from the resin by elution with imidazole solutions. Next, the purified chimeras were bound to an antibody surface by the FLAG sequence, where the salt concentration was reduced, and a high-affinity ligand was introduced to produce a protein-ligand complex. TEV protease and carboxypeptidase A were then introduced to cleave the FLAG and polyhisdidine tags, respectively, from the main part of the ligand-chimera complex. The cleavage products were then purified by size exclusion chromatography.

The chimeric protein-ligand complexes were dissolved in a 10:1 monoolein:cholesterol (Figure 9) mixture and crystals were grown by addition of a solution of consisting of HEPES (for μ and δ complexes only), PEG 400 and a combination of salts (varying between the different chimeras)[30,44,46].

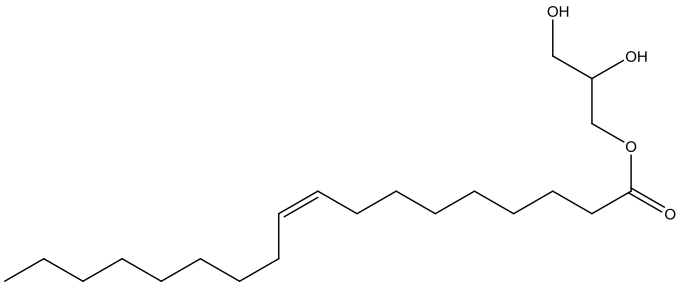



**Figure 9: Structure of monoolein.**

The human κ opioid receptor was used to produce the κ-receptor-T4L chimera. The δ and μ opioid receptor-T4L chimeras were based on the mouse opioid receptors, which are very similar to their human counterparts (Figure 10 and Figure 11)[44]. A sequence alignment of the human and mouse δ receptors (Figure 10) showed that the sequences differ by 23 amino acid residues, 20 of which are in the N and C-terminal regions (94% sequence similarity). The three remaining residues that differ are in the ICLI, ECLII and ECLIII. A sequence alignment of the human and mouse μ receptors also shows a high sequence similarity (94%) (Figure 11). The sequences differ by 25 amino acid residues, 21 of which are in the N and C- terminal regions. The other four residues are in the TMI, TMIII, TMIV and ECLIII.

```
                                (N-Terminus)                                                    (TMI)
sp|P32300|OPRD_MOUSE      MELVPSARAE LQSSPLVNLS DAFPSAFPSA GANASGSPGA RSASSLALAI AITALYSAVC
sp|P41143|OPRD_HUMAN      MEPAPSAGAE LQPPLFANAS DAYPSACPSA GANASGPPGA RSASSLALAI AITALYSAVC
                          ** .*** ** **   :.* * **:*** *** ****** *** ********** **********
                                              (ILCI)          (TMII)                (ECLI)
sp|P32300|OPRD_MOUSE      AVGLLGNVLV MFGIVRYTKL KTATNIYIFN LALADALATS TLPFQSAKYL METWPFGELL
sp|P41143|OPRD_HUMAN      AVGLLGNVLV MFGIVRYTKM KTATNIYIFN LALADALATS TLPFQSAKYL METWPFGELL
                          ********** *********: ********** ********** ********** **********
                                (TMIII)                          (ICLII)        (TMIV)
sp|P32300|OPRD_MOUSE      CKAVLSIDYY NMFTSIFTLT MMSVDRYIAV CHPVKALDFR TPAKAKLIN ICIWVLASGVG
sp|P41143|OPRD_HUMAN      CKAVLSIDYY NMFTSIFTLT MMSVDRYIAV CHPVKALDFR TPAKAKLIN ICIWVLASGVG
                          ********** ********** ********** ********** ********* ***********
                                         (ECLII)                          (TMV)
sp|P32300|OPRD_MOUSE      VPIMVMAVTQ PRDGAVVCML QFPSPSWYWD TVTKICVFLF AFVVPILIIT VCYGLMLLRL
sp|P41143|OPRD_HUMAN      VPIMVMAVTR PRDGAVVCML QFPSPSWYWD TVTKICVFLF AFVVPILIIT VCYGLMLLRL
                          *********: ********** ********** ********** ********** **********
                                (ICLIII)                (TMVI)                (ECLIII)
sp|P32300|OPRD_MOUSE      RSVRLLSGSK EKDRSLRRIT RMVLVVVGAF VVCWAPIHIF VIVWTLVDIN RRDPLVVAAL
sp|P41143|OPRD_HUMAN      RSVRLLSGSK EKDRSLRRIT RMVLVVVGAF VVCWAPIHIF VIVWTLVDID RRDPLVVAAL
                          ********** ********** ********** ********** *********: **********
                                (TMVII)                 (C-Terminus)
sp|P32300|OPRD_MOUSE      HLCIALGYAN SSLNPVLYAF LDENFKRCFR QLCRTPCGRQ EPGSLRRPRQ ATTRERVTAC
sp|P41143|OPRD_HUMAN      HLCIALGYAN SSLNPVLYAF LDENFKRCFR QLCRKPCGRP DPSSFSRARE ATARERVTAC
                          ********** ********** ********** ****.****  :*.*: * *: **:*******
sp|P32300|OPRD_MOUSE      TPSDGPGGGA AA
sp|P41143|OPRD_HUMAN      TPSDGPGGGA AA
                          ********** **
```

**Figure 10: Clustal Omega (1.2.0)[33,34] sequence alignment of the mouse and human WT δ receptors.TMs are shown in bold.**

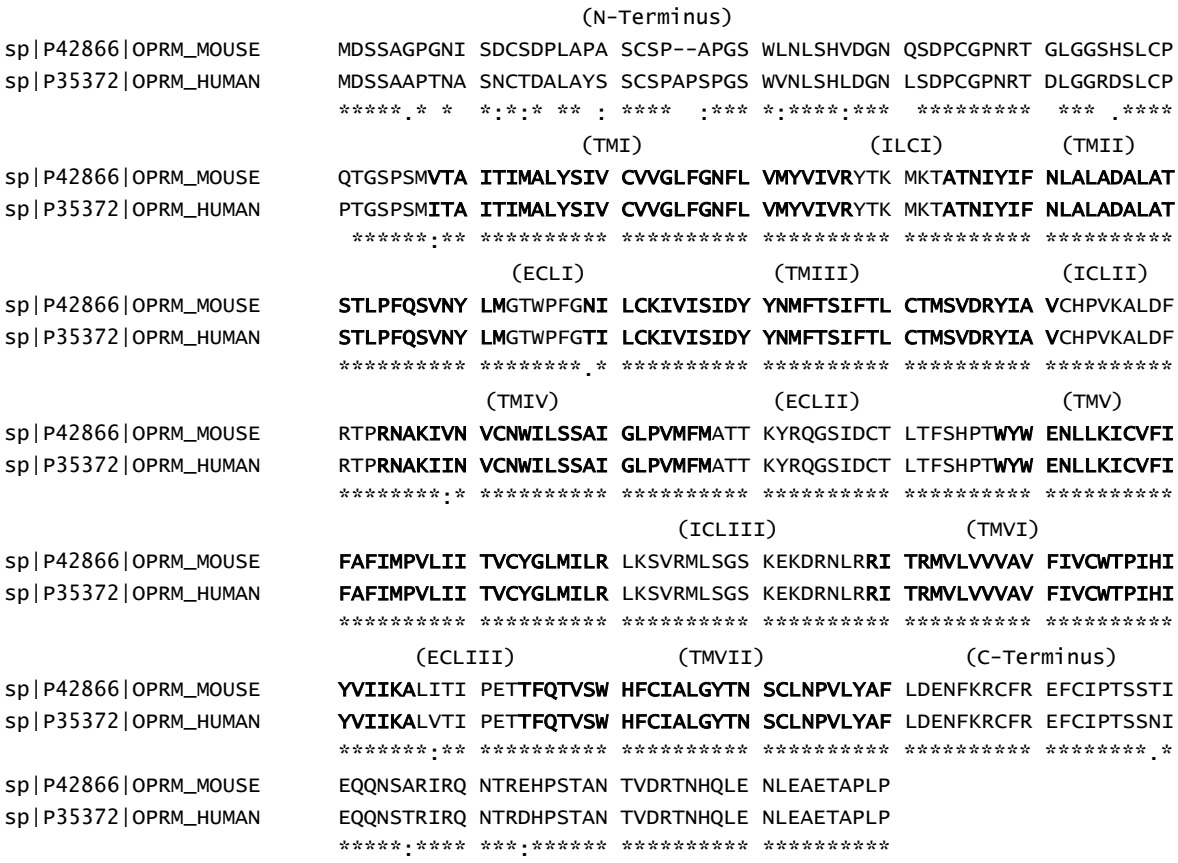



**Figure 11: Clustal Omega (1.2.0)[33,34] sequence alignment of the mouse and human WT μ receptors.TMs are shown in bold.**

The amino acid sequences of the chimeras have two features distinct from that of the native sequences. First, the chimeras have shorter termini (Figure 8) compared to the native receptors (Figure 7). Second, the chimeras have the T4L sequence inserted into the ICLIII (Figure 8).

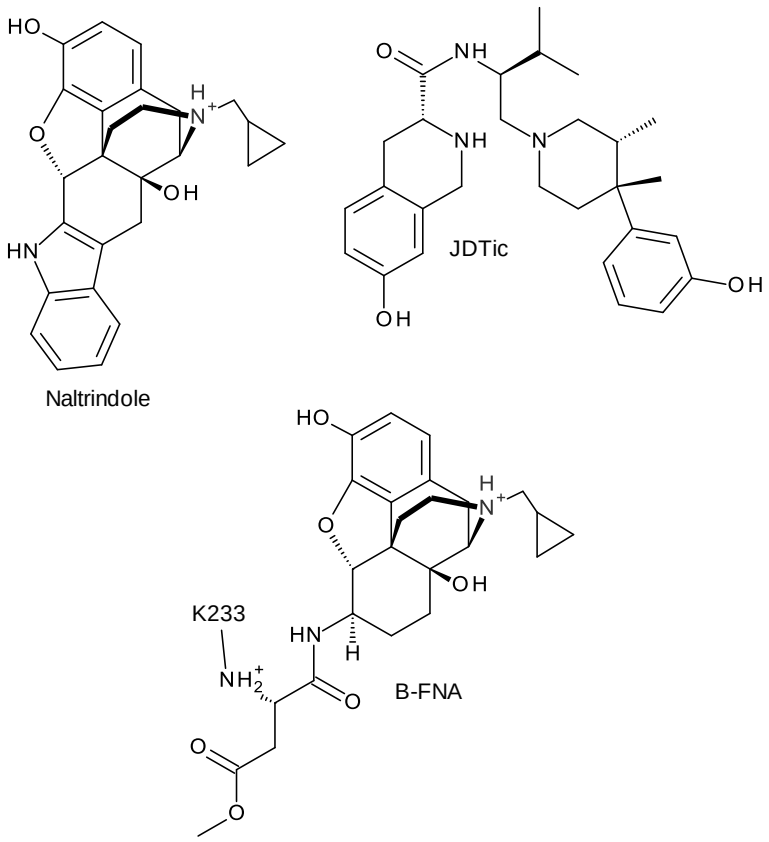



**Figure 12: Structures of the co-crystalized ligands of the opioid receptors. The β-FNA is covalently bound to side chain of K233. The $NH_2^+$ is the amino group of the K233 side chain.**

Each of the receptors was co-crystalized with a ligand (Figure 12) bound to the active site (Figure 13). The ligands bound to the δ, κ and μ receptors were naltrindole[30], (*3R*)-7-hydroxy-N-[(*2S*)-1-[(*3R*,*4R*)-4-(3-hydroxyphenyl)-3,4-dimethylpiperidin-1-yl]-3-methylbutan-2-yl]-1,2,3,4-tetrahydroisoquinoline-3-carboxamide (JDTic)[46] and β-funaltrexamine (β-FNA)[44], respectively.

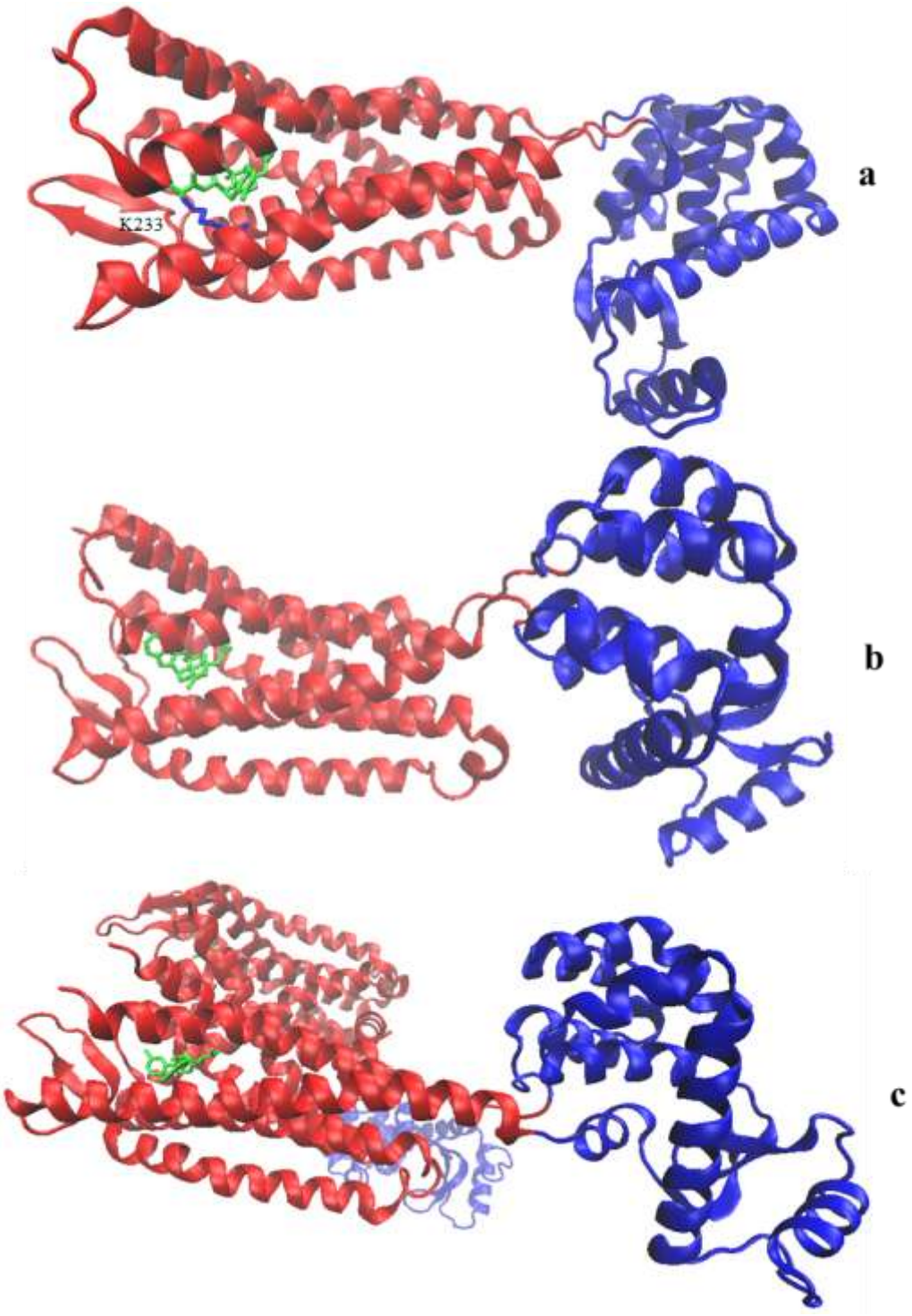



**Figure 13: Crystal structures of a) μ opioid receptor b) δ opioid receptor c) κ opioid receptors. The receptor itself is shown in red, the lysozyme domain is shown in blue and the binding pocket ligand is shown in green.**

In the crystal structure of the μ receptor (Figure 13a), a covalent linkage (blue line) can be seen between the Lys233 side chain and the ligand. The κ crystal structure is a homodimer (Figure 13c).

### 1g: Opioid Receptor Ligands

The endogenous ligands of the opioid receptors, all of which are peptides, fall into several categories. Enkephalins bind preferentially to the δ receptor, dynorphins bind mainly to the κ receptor, while both endorphins and endomorphins bind primarily to the μ receptor[48].

There are many exogenous ligands for the opioid receptors, possessing different selectivity for the three receptors[49]. A selected few are shown in Figure 14 and their inhibition constants for the δ, μ and κ receptors are given in Table 2.

**Table 2: Exogenous ligands and their binding affinities for the δ, μ and κ receptors[49]. The binding affinities were measured based on the displacement of $^3$H labeled Tyr-D-Ala-Gly-Phe-D-Leu (DADLE) in the δ receptor, DAMGO in the μ receptor and U69, 593 in the κ receptor.**

| Compound | $K_I$ ± standard error of the mean (nM) | | |
|---|---|---|---|
| | δ | μ | κ |
| Naltrindole | 1.6 ± 0.1 | 151 ± 14 | 75 ± 7 |
| SNC-80 | 5.6 ± 0.5 | 8070 ± 930 | 8760 ± 710 |
| DAMGO | 469± 39 | 6.1 ± 0.7 | 5820 ± 540 |
| Morphine | 157 ± 11 | 11.0 ± 1.0 | 188 ± 20 |
| U69,593 | > 5000 | 7250 ± 660 | 4.8 ± 0.5 |
| Naltrexone | 7.5 ± 1.8 | 2.4 ± 0.3 | 2.2 ± 0.2 |

**$K_I$=[P][I]/[PI] where [P] is the concentration of free protein (receptor), [I] is the concentration of unbound inhibitor and [PI] is the concentration of protein-inhibitor complex. Therefore a lower $K_I$ means higher binding affinity.**

The six compounds in Table 2 and Figure 14 can be divided into categories based on their structure. Naltrindole, morphine and naltrexone are morphinans; DAMGO is a synthetic peptide; SNC-80 and U69,593 are peptidomimetics.

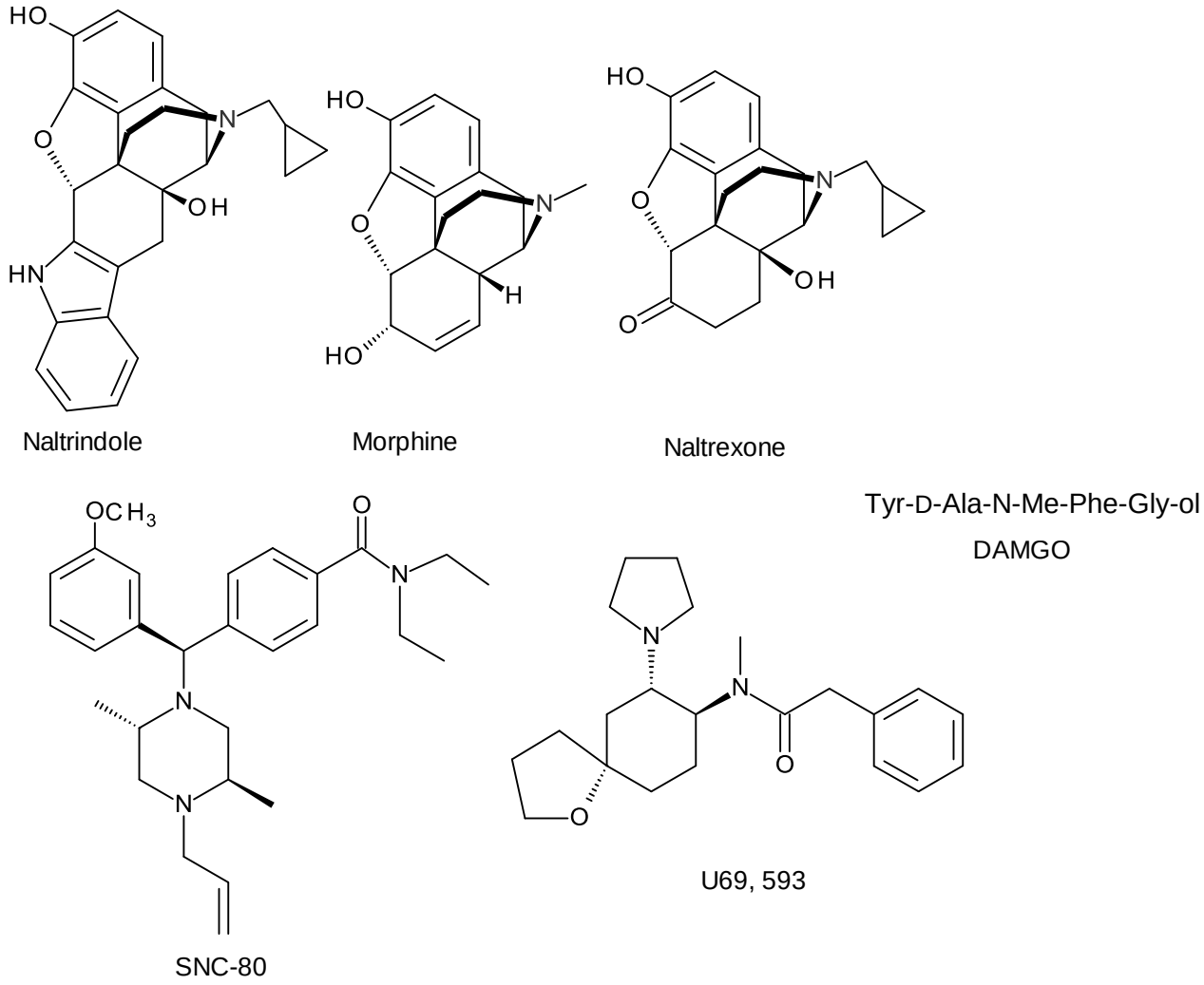



**Figure 14: Structures of the exogenous ligands from Table 2.**

From the $K_i$ values in Table 2, it can be seen that naltrindole and SNC-80 are selective for the δ receptor, while DAMGO and morphine are selective for the μ receptor. U69,593 is κ selective, while naltrexone had no major selectivity for any of the three receptors.

While the highest selectivity is found for peptidomimetic ligands, where the binding affinity for one receptor is over 1000 times higher compared with that for the other two receptors[49], it cannot be assumed that this holds true for all peptidomimetic ligands. This only holds true for the set of ligands shown in Table 2. The same group of researchers that determined the binding affinities in Table 2 also synthesized and measured the binding affinity of a large variety of morphinan analogues. The selectivity of these analogues range from a factor of three to 300 among the three receptors[49]. Another example involved synthetic analogues of enkephalins and endorphins, with differential selectivity of these analogues ranging from a factor of three to 10,000 among the three receptors[50].

***1gi: DAMGO***

DAMGO (Figure 14) has three features that make it unnatural: 1) it contains D-alanine, 2) the phenylalanine is *N*-methylated, and 3) the C-terminal carboxylic acid is reduced to an alcohol. It is worth noting that while D-amino acids are not found in ribosomally-synthesized peptides, they are found in nature. For example, several D-amino acids such as D-alanine, are part of the bacterial peptidoglycan matrix[51]. However, these peptide bonds are made by peptide synthases specific for the oligopeptides of peptidoglycan and not by ribosomes[51].

DAMGO is a completely synthetic peptide that is selective for the μ receptor (Table 2). It is an antagonist that mimics endorphins. Endorphins are one of the classes of endogenous ligands for the μ receptor that possess the sequence YGGFX, where "X" represents a varying C-terminal sequence (see Table 3).

Prior to the publication of crystal structures of the opioid receptors, many binding studies had been performed to determine the receptor amino acids responsible for DAMGO

discrimination among the three opioid receptors[31,32]. The successfully cloning of μ, δ and κ receptors[52-54] in 1993 allowed receptor chimeras and point mutant receptors to be made to study the influence of amino acid substitutions. Three studies are of particular interest[31,32,55].

Minami and co-workers[32] determined that a single point mutation (K108N) in the δ receptor gave the receptor similar binding affinity for DAMGO (μ receptor contains N127 at the same position) as the μ receptor (Figure 7). Furthermore, they determined that the K108 could be substituted by any other amino acid except glycine and tryptophan, to give similar binding affinity for DAMGO as the μ receptor[32].

The second study by Seki and co-workers[31] determined that a four-point mutation was required (E297K, S310V, Y312W and Y313H) to give the κ receptor similar binding affinity for DAMGO to that of the μ receptor (Figure 7). This work involved mutation of four amino acids on the ECLIII of the κ receptor to their μ receptor counterparts. All four mutations were required in order to observe any noticeable improvement in the DAMGO binding affinity of the mutant receptor relative to that of the wild type (WT) κ receptor[31].

Surratt and co-workers[55] showed in their study that the amino acids Asp147 and His297 play a role in DAMGO binding to the μ receptor (Figure 7). When Asp147 was mutated to alanine, glutamate or asparagine, the binding affinity of the mutant μ receptor for DAMGO decreased. When His297 was mutated to alanine, the binding affinity also decreased. The authors proposed that salt bridges form between DAMGO and these two amino acid residues[55]. These two residues are both conserved in all three receptors (μ, κ and δ) (Figure 7).

## 2: Opioid Peptides

### 2a: Endogenous Opioid Receptor Peptide Ligands

A number of endogenous ligands for the opioid receptors are known, all of which are peptide based (Table 3).

**Table 3: Endogenous ligands of the opioid receptors. $–NH_2$ refers to an aminated C-terminus, not the N-terminus**[56].

| Ligand Class (Receptor Selectivity) | Ligand Subtypes | Amino Acid Sequences |
|---|---|---|
| Endorphins | α-Endorphin | YGGFMTSEKSQTPLVT |
| (μ selective) | β-Endorphin | YGGFMTSEKSQTPLVTLFKNAIIKNAYKKGE |
| | γ-Endorphin | YGGFMTSEKSQTPLVTL |
| | α-neo-Endorphin | YGGFLRKYPK |
| | β-neo-Endorphin | YGGFLRKYP |
| Endomorphins | Endomorphin-1 | YPWF-$NH_2$ |
| (μ selective) | Endomorphin-2 | YPFF-$NH_2$ |
| Enkephalins | Leu-Enkephalin | YGGFL |
| (δ selective) | Met-Enkephalin | YGGFM |
| | Met-Enkephalin-Arg-Gly-Leu | YGGFMRGL |
| | Met-Enkephalin-Arg-Phe | YGGFMRF |
| Dynorphins | Dynorphin-A | YGGFLRRIRPKLKWDNQ |
| (κ selective) | Dynorphin-B | YGGFLRRQFKVVT |

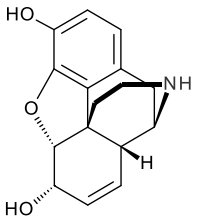
### 2b: Exogenous Opioid Receptor Peptide Ligands

In 1981, Brantl and co-workers showed that fragments of β-casein (one of the proteins in bovine milk) are agonists of the opioid receptor[57], with the highest affinity for the μ- receptor (Table 4). These fragments are known as β-casomorphins, originated from β-casein ("β-caso"),

the parent protein of these peptides, and morphine ("morphin"), which is selective for the μ receptor, just like the β-casomorphins.

**Table 4: $IC_{50}$ (μM) values of exogenous opioid receptor peptide ligand towards guinea pig ileum (GPI) which contains mainly μ receptors and mouse vas deferens (MVD) which contains mainly δ receptors[57].**

| Ligand | Amino Acid Sequence | GPI | MVD |
|---|---|---|---|
| Normorphine | N/A | 0.23 ± 0.022 | 1.1 ± 0.14 |
| Met-Enkephalin | YGGFM | 0.19 ± 0.18 | 0.091 ± 0.002 |
| β-Casomorphin-7 | YPFPGPI | 57.00 ± 7.5 | > 200 |
| β-Casomorphin-6 | YPFPGP | 27.40 ± 1.7 | > 150 |
| β-Casomorphin-5 | YPFPG | 6.50 ± 0.6 | 42.1 ± 5.9 |
| β-Casomorphin-4 | YPFP | 21.90 ± 2.6 | 84.3 ± 12.8 |



Normorphine

The ligands were also tested on rat vas deferens (RVD), which contains mainly κ receptors. No measureable binding affinity was observed for any of the ligands tested, thus no data were displayed for the RVD binding tests (Table 4)[57].

Brantl and co-workers extended their work by performing binding studies on several unnatural analogues of the β-casomorphins. The modifications consisted of substituting the second amino acid (Pro) with D-Ala and aminating the C-terminus of the β-casomorphin analogues (Table 5)[58].

**Table 5: $IC_{50}$ values (μM) to the GPI and MVD. Note that the $–NH_2$ refers to an aminated C-terminus and not the N-terminus.[58] The shaded entries have a selectivity ≥5.0.**

| Ligand | GPI (μ) | MVD (δ) |
|---|---|---|
| Tyr-Pro-Phe-Pro | 22 | 84 |
| Tyr-Pro-Phe-Pro-$NH_2$ | 1.45 | 2.0 |
| Tyr-D-Ala-Phe-Pro-$NH_2$ | 0.40 | 1.0 |
| Tyr-Pro-Phe-Pro-Gly | 6.5 | 40 |
| Tyr-Pro-Phe-Pro-Gly-$NH_2$ | 3.5 | 7.0 |
| Tyr-D-Ala-Phe-Pro-Gly-$NH_2$ | 5.0 | 8.5 |
| Tyr-D-Ala-Phe-Pro-Gly | 2.3 | 3.0 |
| Tyr-D-Ala-Phe-Pro-Met-$NH_2$ | 3.0 | 0.70 |
| Tyr-D-Ala-Phe-Pro-Tyr-$NH_2$ | 0.10 | 0.20 |
| Tyr-Pro-Phe-Pro-Gly-Pro | 27.4 | > 150 |
| Tyr-Pro-Phe-Pro-Gly-Pro-$NH_2$ | 8.0 | 40 |
| Normorphine | 0.20 | 1.1 |
| Met-Enkephalin | 0.19 | 0.019 |

It is important to note that the unmodified β-casomorphins have the highest selectivity for the μ receptor over the δ receptor (Table 5: row 4 and 10). Inclusion of D-Ala and C-terminal amidation do increase the binding affinity for the receptor (Table 5: row 2-3, 5-10); however, the binding affinity for the δ receptor increases by a greater proportion. Thus, these analogues have lower selectivity for the μ receptor. No notable binding affinity was measured for any of the ligands when tested against RVD and therefore, these ligands do not bind to the κ receptor[58].

In both publications by Brantl and co-workers[57,58], the binding affinity was not measured directly. Rather, the tissue was treated with the ligand and then subjected to an electrical pulse treatment[57,58]. It has been shown that tissues containing GPCRs will undergo electrically stimulated contractions. The presence of an agonist in the GPCR has inhibitory effects on the extent of this contraction[12]. Thus, the binding affinity of the ligands towards the opioid receptors can be measured in an indirect manner.

In 1985, Koch and co-workers determined the binding affinity of β-casomorphin-4,5,7 and 8 using homogenized rat brains[59]. A competitive binding experiment was done where the homogenate was pre-treated with either [$^3$H]DAMGO, [$^3$H]DADLE or [$^3$H]ethylketocyclazocine. $IC_{50}$ of the β-casomorphin was determined based on the amount of the radiolabeled ligand displaced (Table 6). The extent of displacement of the radiolabelled ligand with known selectivity indicates the opioid receptor for which β-casomorphin has the greatest affinity.

**Table 6: A list of the ligands with their $IC_{50}$ values (μM) to homogenized rat brain. DADLE: Tyr-D-Ala-Gly-Phe-D-Leu, a δ selective agonist.**

| Ligand | Amino Acid Sequence | [$^3$H]DAMGO μ-selective | [$^3$H]DADLE δ-selective | [$^3$H]Ethylketocyclazocine κ-selective |
|---|---|---|---|---|
| β-Casomorphin-8 | YPFPGPIP | 0.8 | 17 | 60 |
| β-Casomorphin-7 | YPFPGPI | 2 | 12 | 340 |
| β-Casomorphin-5 | YPFPG | 0.4 | 43 | 91 |
| β-Casomorphin-4 | YPFP | 1.4 | 51 | 1800 |

The lower $IC_{50}$ values of β-casomorphins for [$^3$H]DAMGO versus [$^3$H]DADLE and [$^3$H]ethylketocylazocine indicated that the β-casomorphins are selective for the μ receptor, which was consistent with Brantl's observations[57](Table 5). It can also be seen that β-casomorphin-5 displayed the highest binding affinity for the μ receptor[59], which is also consistent with previous finding by Brantl and co-workers.

Subsequent studies on β-casomorphin binding were performed on analogues of the peptides[60]. Many of the modified sequences tested involved one or more D-amimo acid substitutions, usage of unnatural amino acids (e.g. pipecolic acid)[60] and/or cyclic peptides[61].

While the binding affinities have been determined for these β-casomorphins and their analogues, there is no information on which non-conserved amino acids in the μ, δ and κ receptors account for the selectivity. Neither is there information on the conformation of these β-

casomorphins when they bind to the μ receptor. The goal of the present study is to provide insight by predicting the conformation of these β-casomorphins in the μ receptor by using computational molecular docking methods. In doing so, the non-conserved amino acids in the μ, δ and κ receptors that are responsible for the selectivity of β-casomorphins can be predicted. This study will focus on the unmodified β-casomorphins 5 and 7.

## 3: Computational Molecular Docking

### 3a: Genetic Algorithm

Computational molecular docking involves the use of an algorithm to locate the conformation at which the energy of the interaction of two molecules (for example a ligand and a protein) is optimal[62]. The first docking program, DOCK, was developed by Kuntz and co-workers in 1982[63].The algorithm this program had only allowed for rigid ligand docking, meaning that the algorithm was able to manipulate the location of the centre-of-mass of the ligand in space and orientation (solid body rotation). It could not modify the internal conformation of the ligand (thus called rigid docking)[63]. Leach and Kuntz improved the DOCK program in 1992 by upgrading the algorithm to allow flexibility of the ligand. That is, in addition to the algorithm being able to manipulate the centre-of-mass and the orientation of the ligand, it could also modify the conformation (dihedral angles) of the ligand[64]. In 1994, Leach extended this algorithm to allow flexibility in the protein side chains[62]. Therefore, in addition to allowing the conformation of the ligand to be modified, the conformation of the protein (receptor) side chains can also be modified. Genetic algorithms (GAs) were first reported to be used for docking ligands in proteins by Oshiro, Dixon and Kuntz in 1995[65]. GAs were much faster in docking flexible ligands in comparison to the previous algorithms that were available[65]. In the same year,

Jones and co-workers published an extension of the GAs to permit flexibility in the side chains of the protein[66]. DOCK[65], FlipDock[67], and GOLD[68] are examples of docking programs that use GAs.

The GA algorithm works with a population of "chromosomes" that encode information pertaining to a molecular conformation. These chromosomes are composed of a set of "genes" (in reference to the biological counterpart). Each of the rotatable dihedral bond angles of the ligand has its own "gene"[66]. In addition there are two additional genes present - one to represent the Cartesian coordinates of the centre-of-mass of the ligand and the other to represent the orientation (solid body rotation) of the ligand in the form of Euler angles (Figure 15c). Furthermore, if side chains of the protein are allowed to be flexible, then each rotatable bond of those side chains will be described by a "gene". Figure 15a describes the chromosome representation of butane where Figure 15b indicates the dihedral angle.

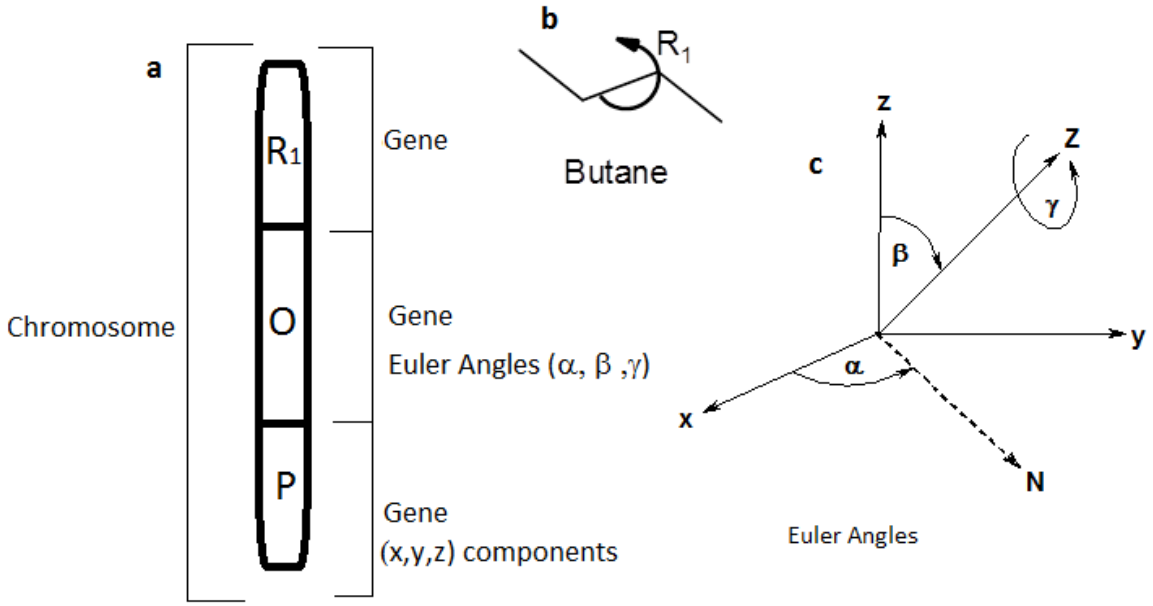



**Figure 15: Schematic representation of the "chromosome" of butane and its "genes". a) The chromosome is defined as the entire ellipse composed of the three genes "$R_1$", "O" and "P". "$R_1$" is the gene that encodes for the dihedral angle of the central C-C bond of the butane. Gene "O" encodes for the orientation of butane and gene "P" encodes for the centre-of-mass of butane. b) Structure of butane with an arrow indicating the rotation of the central C-C bond. c) Definition of the Euler angle where x, y and z are the reference axes and X, Y and Z are the axes of rotatable object (e.g. ligand) which are fixed to the object. N is the node between the xy and XY planes. α is the angle between the x axis and N, γ is the angle between the X axis and N and β is the angle between the z and Z axes.**

The GA algorithm consists of the following steps. An initial population of chromosomes is generated with all the values represented by the constituent "genes" chosen at random[66]. Each chromosome represents a random conformation of the ligand in a random position and orientation within the protein binding site. New and "fitter" populations are generated throughout computational crossovers and mutations and selection events. A crossover occurs between randomly chosen chromosome pairs in which a random number of genes are exchanged between the chromosomes in a single crossover.

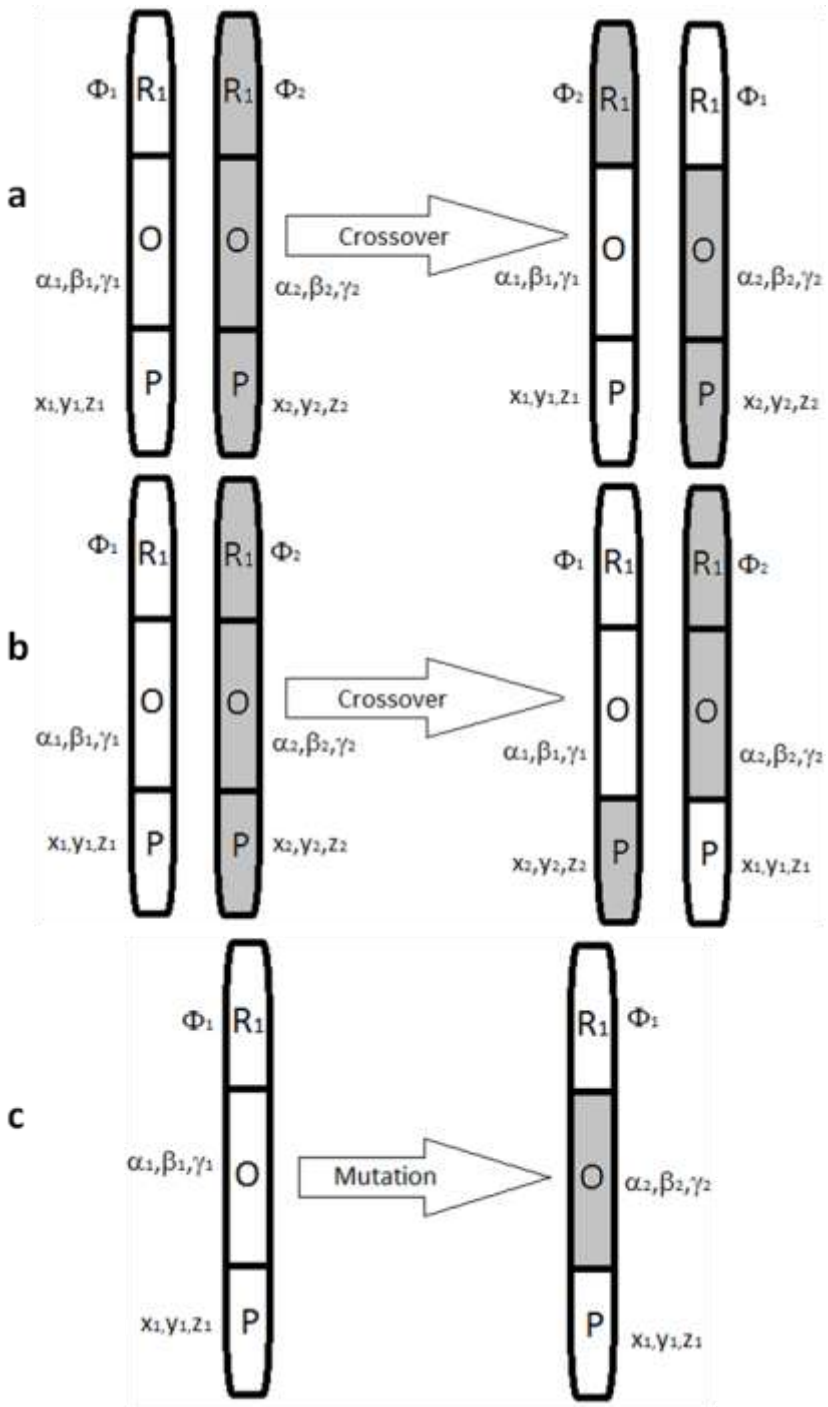



**Figure 16: Schematic representation of crossovers and mutations. "$R_1$" is the gene that encodes for the dihedral angle of the central C-C bond (Φ) of the butane molecule. Gene "O" encodes for the orientation of the butane (α,β,γ) and gene "P" encodes for the centre-of-mass of butane in space (x,y,z). a) A crossover where the dihedral angle gene ("$R_1$") is exchanged between two chromosomes, b) a crossover where the centre-of-mass gene ("P") is exchanged, and c) a mutation of the orientation gene ("O") causing a random re-orientation of the molecule in 3-dimensional space.**

Figure 16 illustrates a crossover between two conformations of butane using the "chromosome" depicted in Figure 15a. When the gene encoding for the dihedral angle ("$R_1$") undergoes a crossover, the values of the dihedral angles of two different conformations are exchanged. However, this has no influence on the centre-of-mass or orientation of either butane model (Figure 16a). When the gene encoding for the centre-of-mass ("P") undergoes a crossover, the centres-of-mass of the two butane models are exchanged. However, this has no influence on the dihedral angle or molecular orientation either (Figure 16b). In addition, mutations are programmed, in which a random gene will be changed randomly. Using the butane example again, if the gene encoding for the orientation "O" underwent a random mutation, the butane model would adopt a new random solid body orientation (Figure 16c). As with the crossovers, this single mutation does not affect the dihedral angle or centre-of-mass of butane. A notable difference between crossovers and mutations is that crossovers require two chromosomes while mutations only require one chromosome[66]. The relative frequency with which crossovers and mutations occur in a GA is selected by the user.

The goal of the GA process is to generate chromosomes with better fitness scores (described in section 3c). The new chromosomes (from both crossovers and mutations) are rescored along with all the chromosomes of the original population. The worst scoring chromosomes of this population are replaced by the new chromosomes. This process is repeated many times (typically thousands of times). The chromosome score is a measure of their fitness. The score improves when favourable interactions are made between the ligand and the protein. The score worsens when steric and repulsive interactions are present. Details on scoring functions will be discussed in section 3c.

The number of chromosomes may be held constant throughout the docking experiment. The exception to this is when "elitism" is programmed into the GA. If the number of chromosomes is held constant at "n", then elitism will results in there being n + 1 chromosome. Elitism occurs when a chromosome is detected with a very good score. This chromosome acquires elitism status, which will allow it to engage in crossovers but not in mutations. If a higher scoring chromosome is generated, then it will replace the previous chromosome with the elitism status. The purpose of elitism is to prevent very good scoring chromosomes from being lost to subsequent crossovers and mutations[69].

It is not always possible to find the optimal ligand-protein interaction using GA. The reason for this is that when a ligand has many rotatable bonds (each encoded by a separate gene), the number of possible crossover sites will be large. Thus, there are many possible chromosomes than can be formed from a single pair of parent chromosomes. There will be chromosomes formed by the GA with a good score (but not necessarily the best score). Once this occurs, the ability for the GA to further evolve the chromosomes is reduced. The reason is that generating new chromosomes that have a better score than a high scoring parent is unlikely when the number of crossover/mutation sites is high[69].

**3b: Genetic Optimization for Ligand Docking (GOLD)**

Several programs are available to bind (dock) ligands to the proteins, such as DOCK[70], FlexX[71], Surflex[72] and GOLD[73]. All of these programs allow for the internal flexibility of the ligand. Each program has been able to accurately predict the binding conformation of ligands in their known receptors. Studies have been done comparing differences between docking programs (including DOCK, Surflex and GOLD)[74,75]. The conclusion was that there is no single docking program that outperforms all the other programs for every class of ligand and receptor. As a

consequence, GOLD has been chosen as the docking program for this study because it has an easy interface to use, and it performed well in some docking comparisons[73].

GOLD, like other docking programs, allows the parameters of the docking experiment to be set. These parameters include selecting the binding pocket, choosing the scoring function, introducing protein side chain flexibility and introducing covalent attachments of the ligand to the protein. Binding pocket selection can be done in one of two ways. First, one can indicate all the amino acid residues that constitute the binding pocket. Second, one can indicate an atom in a central location in the binding and indicate a specific radius from that atom as the binding pocket.

The GA of GOLD has several features in addition to what was described in Section: 3a[76]. The first is the presence of islands. What had been observed from GAs is that the most fit chromosomes consistently engage in crossovers while low scoring chromosomes rarely engage in crossovers. This results in the same few chromosomes participating in most of the crossovers which in turn results in less diversity among the chromosomes. Islands consist on populations of chromosomes that can only crossovers within their own island. Even when highly fit chromosomes form on each island, the multiple islands mean that there are multiple sets of such chromosomes which each have evolved independently, therefore resulting in an overall increase in chromosome diversity compared to having a single population. Migration involves the migration of one randomly selected chromosome from one island to another. This is coupled with a randomly selection chromosome from the second island migration to the first island in order to maintain a contain population on each island. This mechanism is in place to further enhance the diversity among the chromosomes. The final feature in the GA of GOLD is call a niche. The purpose of niches is to prevent the accumulation of chromosomes representing

geometrically similar poses in order to further diversify the chromosomes. Chromosomes are determined to belong to the same niche if the root mean square deviation (RMSD, refer to Section: 3d) of the H-bonding heteroatoms is less than 1.0Å. Once a niche is filled and a new chromosomes of that niche is generated, then the lowest scoring chromosome of that niche is eliminated[76].

**3c: Scoring Function**

As was mentioned in section 3a, the docked poses of ligands in the receptor site generated by the GA are scored for their fitness. Various scoring functions are available for this purpose. Some examples of scoring functions available in GOLD are GOLDScore, CHEMScore, ASP and CHEMPLP[77]. The functions CHEMScore and CHEMPLP are described in section 3ci and 3cii. These two scoring functions were used in this study.

***3ci: CHEMScore***

CHEMScore was developed by Eldridge and co-workers in 1997 for use in the docking program PRO_SELECT[78]. This scoring function calculates the free energy ΔG of the ligand binding to the protein:

$$\Delta G = \Delta G^{\circ} + \Delta G_{hbond} \sum_{iI} g_1(\Delta r)\, g_2(\Delta \alpha) + \Delta G_{metal} \sum_{aM} f(r_{aM}) + \Delta G_{lipo} \sum_{lL} f(r_{lL}) + \Delta G_{rot} \Delta H_{rot} \quad \textbf{(1)}$$

The $\Delta G_{hbond}$ term determines the contribution from H-bonding. The variable Δr describes the deviation from the optimal H-bonding distance of 1.85Å for each pair of H···O/N atoms. The variable Δα describes the deviation from the optimal angle of 180° for each N/O-H···O/N pair. The indices i and I represent the atoms of the ligand and protein, respectively, participating in an H-bond.

In equation (1), the $\Delta G_{metal}$ term determines the contribution to the free energy of binding from metal coordination. The variable $r_{aM}$ is the distance between the metal centre and coordinating atom. The indices a and M represent the coordinating atoms of the ligand and metal centre of the protein, respectively.

The $\Delta G_{lipo}$ term in equation (1) determines the contribution to the free energy of binding from hydrophobic interactions. The variable $r_{lL}$ is the distance between a hydrophobic atom "l" on the ligand and a hydrophobic atom "L" on the protein. The terms $\Delta G°$, $\Delta G_{hbond}$, $\Delta G_{metal}$ and $\Delta G_{rot}$ are constants.

The last term in equation (1) determines the penalty from the loss of rotation of the rotatable bonds upon binding to the ligand. The $\Delta H_{rot}$ term is calculated as:

$$\Delta H_{rot} = 1 + \left(1 - \frac{1}{N_{rot}}\right) \sum_{i} \frac{\left(P_{nl}(i) + P'_{nl}(i)\right)}{2} \qquad \textbf{(2)}$$

In equation (2), $N_{rot}$ is the number of rotatable bonds in the ligand that must be frozen in order to make an interaction with the protein. Groups of the ligand that form interactions with the protein will have their dihedral angles frozen in order to make the interaction. Groups of the ligand that do not make any interactions with the protein will still be able rotate their dihedral angles. The index i represents the rotatable bonds. The two terms $P_{nl}$ and $P'_{nl}$ represent the percentages of all the atoms (excluding hydrogen atoms) on either side of the rotatable bond that are non-lipophilic. Atoms defined as lipophilic are the following: halogens, sulfur with no additional heteroatoms or hydrogen atoms bound to it, and any carbon atom except those of carbonyl, nitrile, and any group containing two or more heteroatoms on the carbon atom. Thus any other functional group is considered non-lipophilic[78].

### *3cii: Piecewise Linear Potential (CHEMPLP)*

CHEMPLP was developed in 2009 by Korb and co-workers[68] and is the most recent scoring function implemented in GOLD. This function uses the proximity of atom pairs between the ligand and receptor to deduce a score for the interaction of those two atoms. For CHEMPLP, large positive scores indicate good ligand-receptor interactions.

$$\begin{aligned} \text{CHEMPLP} &= f_{\text{CHEMPLP}} \\ &= f_{\text{PLP}} + f_{\text{hb}} + f_{\text{hb-ch}} + f_{\text{hb-CHO}} + f_{\text{met}} + f_{\text{met-coord}} + f_{\text{met-ch}} \\ &\quad + f_{\text{met-coord-ch}} + f_{\text{clash}} + f_{\text{tors}} \end{aligned} \qquad \textbf{(3)}$$

Each of the ten terms in equation (3) will be discussed in turn.

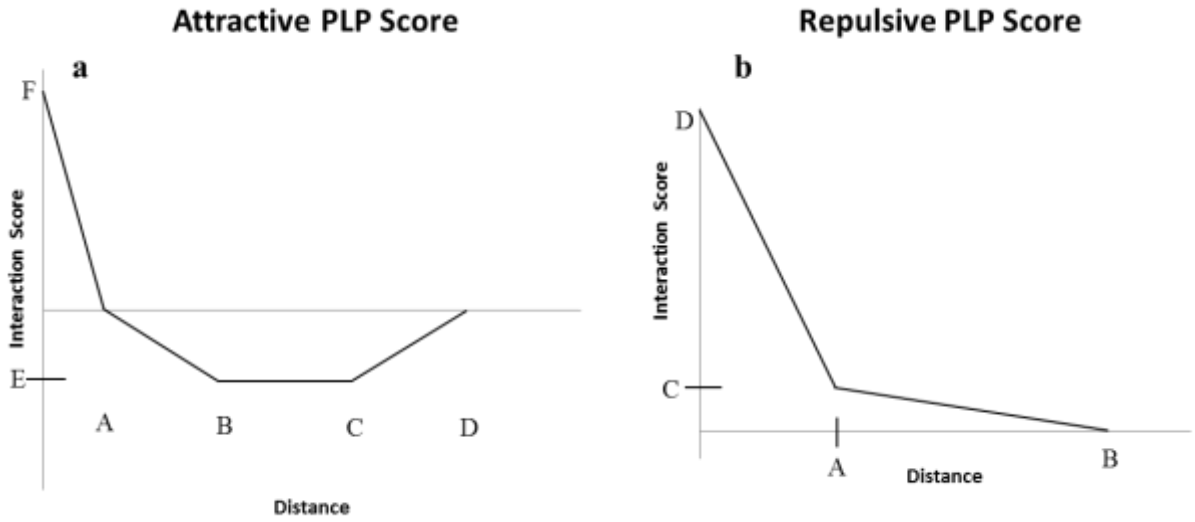



**Figure 17: Relationship between score of an interaction and distance between interacting atoms of (a) attractive and (b) repulsive atom pairs.**

The term $f_{PLP}$ in equation (3) describes the score contribution from the proximity (scalar distance) of protein-ligand atom pairs. The equation for this calculation is comprised of two terms, one for the attractive contribution and one for the repulsive contribution:

$$f_{PLP} = \sum_{p \in P_{prot-lig-plp}} plp(p_r, p_A, p_B, p_C, p_D, p_E, p_F) + \sum_{p \in P_{prot-lig-rep}} rep(p_r, p_A, p_B, p_C, p_D) \quad \textbf{(4)}$$

The first term describes the attractive interaction between a protein-ligand atom pair where the plp is a piecewise linear potential (PLP) function, which is the sum of four linear functions shown in Figure 17a. The distance between the pair of atoms is the variable $p_r$ and the constants $p_A$ to $p_F$ refer to the six points (A to F) in Figure 17a. In this graph/function, a more negative score is more beneficial to the overall fitness score and a more positive score in the graph is detrimental to the overall score. The x-axis in Figure 17a represents the distance $p_r$ between the two atoms. If the distance between the atoms is greater than D, then interaction $plp(p_r)$ will have a score of 0. Between B and C, the distance between the atoms is optimal i.e. the shortest distance without steric clashing. At distances below A, the atoms overlap, thus the score increases. The use of a PLP rather than a continuous function makes calculation of the docking score very efficient.

The second term in equation (4) describes the repulsive interaction between a protein-ligand atom pair as shown in Figure 17b. The distance between the atom pair is the variable $p_r$ and the constants $p_A$ to $p_D$ refer to the four points (A to D) in the graph. The x-axis represents distance $p_r$ between the two atoms. If the distance $p_r$ between the atoms is greater than B, then the interaction $rep(p_r)$ will have a score of 0. If the distance between the atoms is between points A and B, then the linear function describes dipole-dipole repulsion. If the distance is less than A, however, then the atoms experience both dipole-dipole and steric repulsion.

The only atom pairs for which a repulsive PLP score is evaluated are the following: H-bond donor:H-bond donor, H-bond acceptor:H-bond acceptor and H-bond donor:metal centre. Each of these atom combinations have repulsive interactions. All other atom combinations are

described by the attractive PLP described in Figure 17a, including interactions between polar and non-polar atoms, which are designated "buried" in the scoring function. The values of constants A to F vary depending on the type of interaction formed between the atoms. For example, an H-bond, buried interaction, and hydrophobic interaction all fall under the attractive score contributions (Figure 17a). However the magnitude of the interaction will be different for each of these interaction types.

In equation (3), the terms $f_{hb}$, $f_{hb\text{-}ch}$ and $f_{hb\text{-}CHO}$ refer to H-bonding contributions to the score. The nature of the H-bond determines which of these functions are used. The $f_{hb}$ function is used when either one or both of the H-bonding functional groups are neutral (uncharged) (Figure 18a). The $f_{hb\text{-}ch}$ function is used when both of the H-bonding groups are charged (Figure 18b). The $f_{hb\text{-}CHO}$ is used when a C-H group H-bonds with an oxygen. The only C-H groups that are permitted to do this are aromatic C-H groups *ortho* to aromatic (cyclic) nitrogens (Figure 18c). The common place to see such a C-H is in the amino acid tryptophan. The 2-position C-H of the indole ring meets the criterion to be an H-bond donor.

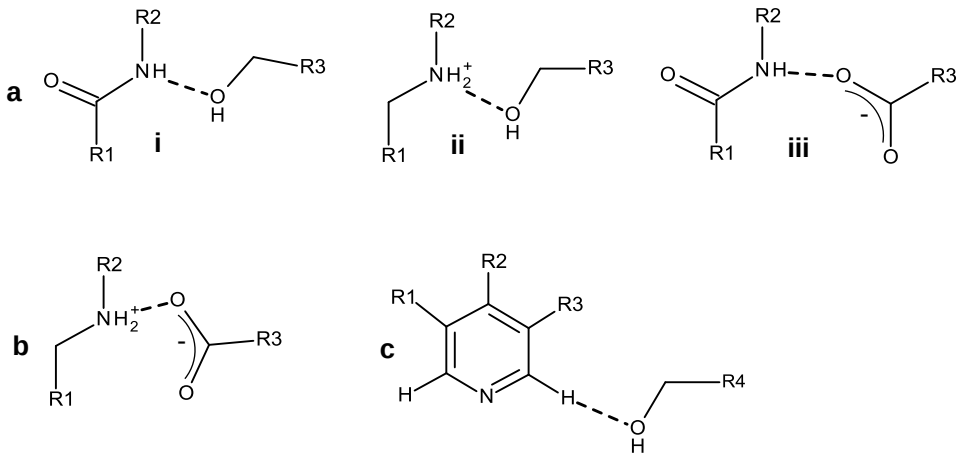



**Figure 18: H-bonding (dash lines) classification used by CHEMPLP a) examples of H-bonds that utilize the $f_{hb}$ function (equation (5)), i) two neutral molecules H-bonding, ii) a cationic and neutral molecule H-bonding, and iii) an anionic and neutral molecule H-bonding; b) an example of an H-bond that utilizes the $f_{hb\text{-}ch}$ function (equation (6))-the formation of a salt bridge; and d) an example of an H-bond that utilizes the $f_{hb\text{-}CHO}$ function (equation (7))-a CH *ortho* to an aromatic nitrogen H-bonded to an oxygen.**

The PLP functions used to calculate the three hydrogen bonding contributions to the CHEMPLP score are:

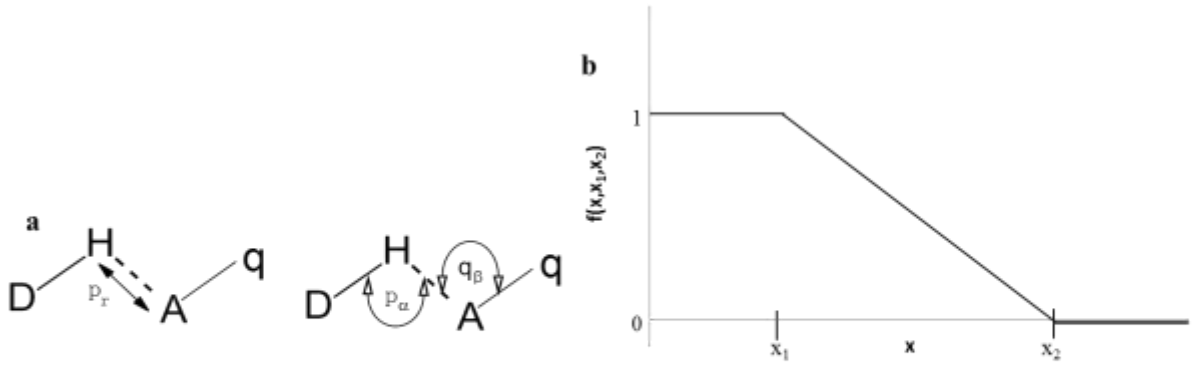
$$f_{hb} = w_{hb} \cdot \sum_{p \in P_{hb}} f(|p_r - 1.85|, 0.25, 0.65) \cdot f(|p_\alpha - 180|, 30, 80) \cdot \prod_{q \in P_{acc_{nb}}} f\left(|q_\beta - 180|, 80, 100\right) \quad \textbf{(5)}$$

$$f_{hb-ch} = w_{hb} \cdot w_{hb-ch} \cdot \sum_{p \in P_{hb-ch}} f(|p_r - 1.85|, 0.25, 0.65) \cdot f(|p_\alpha - 180|, 30, 80) \cdot \prod_{q \in P_{acc\_nb}} f\left(|q_\beta - 180|, 80, 100\right) \quad \textbf{(6)}$$

$$f_{hb-CHO} = w_{hb-CHO} \cdot \sum_{p \in P_{hb-CHO}} f(|p_r - 2.35|, 0.25, 0.65) \cdot f(|p_\alpha - 180|, 50, 100) \cdot \prod_{q \in P_{acc_{nb}}} f\left(|q_\beta - 180|, 80, 100\right) \quad \textbf{(7)}$$

Equation (5) describes H-bonding between two neutral groups or one neutral and one charged group. Equation (6) describes H-bonding between two charged groups. Equation (7) describes H-bonding between an aromatic C-H groups *ortho* to an aromatic (cyclic) nitrogen and an oxygen.

In equations (5)-(7), $w_{hb}$, $w_{hb\text{-}ch}$ and $w_{hb\text{-}CHO}$ are weighting factors. The variable $p_r$ describes the distance between the hydrogen atom and H-bond acceptor. The two variables $p_\alpha$ and $q_\beta$ represent three-atom angles DHA and HAq, respectively, formed by the H-bond (Figure 19a).




$$f(x,x_1,x_2)=\begin{cases}1 & if\ x\le x_1\\ (x_2-x)/(x_2-x_1) & if\ x_1<x\le x_2\\ 0 & if\ x>x_2\end{cases}$$

**Figure 19: Definition of $p_r$, the $p_\alpha$ and $q_\beta$ angles and the constants for H-bonding functions in CHEMPLP. a) "H", "D", "A" and "q" represent the H-bonding hydrogen, the atom covalently bound to the H-bonding hydrogen, the H-bond acceptor and the atom covalently bound to the H-bond acceptor, respectively; b) A PLP function used in equations (5)-(7).**

A set of PLP functions are used to describe the influence of the three variables: $p_r$, $p_\alpha$ and $q_\beta$ (Figure 19b). Consider the function $f(|p_r-1.85|,0.25,0.65)$ in equation (5). The ideal H-bond length is 1.85Å. Therefore $|p_r-1.85|$ is the deviation from the ideal H-bond length

(equation (5)) and is represented by the variable "x" in the PLP function depicted in Figure 19b. The two subsequent constants in this function, 0.25 and 0.65, represent $x_1$ and $x_2$ respectively. If the deviation of the H-bond from the ideal length is less than 0.25Å, then the PLP function will be assigned a value of 1. If the deviation of the H-bond from the ideal length is greater than 0.65Å, then the PLP function will be assigned a value of 0. If the deviation of the H-bond from the ideal length is between 0.25Å and 0.65Å, then the PLP function will follow the linear equation $f(x, x_1, x_2) = (0.65 - x)/0.4$. The functions $f(p_\alpha)$ and $f(q_\beta)$ of equation (5) work analogously.

The next four terms in the CHEMPLP score (equation (3)) are $f_{met}$, $f_{met\text{-}coord}$, $f_{met\text{-}ch}$ and $f_{met\text{-}coord\text{-}ch}$. All these terms pertain to the contribution from the ligand coordinating to a metal centre on the protein (equations (8) and (9)):

$$f_{met} = w_{met} \cdot \sum_{p \in P_{acc-met}} f(p_r, 2.6, 3.0) \cdot \prod_{q \in P_{acc\_nb}} f(|q_\beta - 180|, 80, 90) \quad \textbf{(8)}$$

$$f_{met-coord} = w_{met} \cdot \sum_{p \in P_{acc-met-coord}} f(p_{rf}, 0.6, 0.8) \cdot \prod_{q \in P_{acc\_nb}} f(|q_\beta - 180|, 80, 90) \quad \textbf{(9)}$$

Similar to the H-bonding equations, $w_{met}$ is a weighting factor, $p_r$ is the distance between the metal centre and coordinating atom A of the ligand and $q_\beta$ is the MAq angle (Figure 20a).

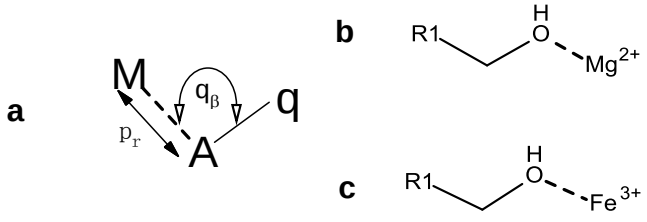



**Figure 20: Metal coordination (dash lines) classification used by CHEMPLP a) Definition of $p_r$ and the $q_\beta$ angle. "M" is the metal centre, "A" is the coordinating atom and "q" is the atom covalently bound to the coordinating atom, b) an example of a coordination that uses equation (8), and c) an example of a coordination that uses equation (9).**

Equations (8) and (9) are PLP functions similar to the functions used for H-bonding (equations (5)-(7)).

The next term in equation (3) is $f_{clash,}$ which describes steric clashes within the ligand. This term is not to be confused with protein-ligand steric clashes. While the latter is part of the $f_{PLP}$ function as described earlier (equation (3)), the $f_{clash}$ term deals with steric clashes that occur between atoms of the ligand itself:

$$f_{clash} = w_{clash} \cdot \sum_{c \in C_{clash}} f_{dist}(\left\| \overrightarrow{C_{p_\alpha} C_{p_b}} \right\|, C_w, C_{r_{clash}}) \quad \textbf{(10)}$$

The variable $w_{clash}$ is the weighting factor for this function, $\overrightarrow{C_{p_\alpha} C_{p_b}}$ is the vector between the clashing atoms α and b of the ligand, $C_w$ is the weighting factor of the $f_{dist}$ function and $C_{rclash}$ is the clash cut-off distance. If the distance between the atoms as determined by $\left\| \overrightarrow{C_{p_\alpha} C_{p_b}} \right\|$ is greater than $C_{rclash}$, then $f_{clash}$=0 (Figure 21a) for that atom set. If the distance between the atoms is less or equal to $C_{rclash}$, then $f_{clash}$>0, which is detrimental to the score (Figure 21b).

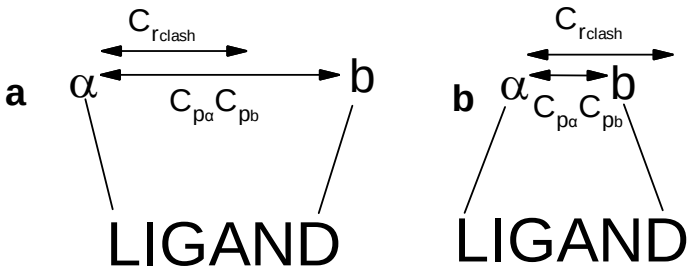



**Figure 21: Schematic of the clash cut-off where "α" and "b" are two atoms of a ligand and "LIGAND" is all the atoms connecting atoms "α" and "b". a) ligand conformation where the distance between "α" and "b" is greater than the clash cut-off and b) ligand conformation where the distance between "α" and "b" is less than the clash cut-off.**

The final term in CHEMPLP, equation (3), is $f_{tors}$, which describes the score penalty for the loss of bond rotation from the rotatable bonds of the ligand:

$$f_{tors} = w_{tors} \sum_{p \in P_{tors}} \frac{1}{2} p_V (1 + p_S \cos(|p_n| p_\omega)) \quad \textbf{(11)}$$

The constant $w_{tors}$ in equation (11) is the weighting factor for this function, $p_V$ is the torsional barrier (energy required for rotating the bond), and $p_s$ determines whether the minimum energy conformation is staggered or eclipsed. The $p_n$ represents the periodicity of the function and the variable $p_\omega$ is the value of the rotatable dihedral angle[68].

**3d: Root Mean Squared Deviation**

Root mean squared deviation (distance) (RMSD) is a method to measure how similar two structures are to each other. It is a valuable tool in docking studies as it provides a qualitative comparison of a docked pose to that of the crystal structure (if available). The term "pose" is defined as the combination of the conformation of the ligand, orientation and position (in three-dimensional space) of the ligand relative to the protein. The RMSD allows the quality of docked poses of a ligand to be measured quantitatively (equation (12)).

$$RMSD = \sqrt{\sum_{i=1}^{n} \frac{\left((x_{i_2} - x_{i_1})^2 + (y_{i_2} - y_{i_1})^2 + (z_{i_2} - z_{i_1})^2\right)}{n}} \quad \textbf{(12)}$$

The variables ($x_{i2}$,$y_{i2}$,$z_{i2}$) and ($x_{i1}$,$y_{i1}$,$z_{i1}$) are the coordinates of the $i^{th}$ atom of the docked and reference (e.g. crystal structure) pose, respectively. The "n" refers to the total number of atoms in the ligand. This equation calculates the extent to which the atoms of the docked pose deviate from the crystal structure. Therefore the lower the RMSD value, the more similar the docked pose is to the crystal structure.

Since no hydrogen atoms are present in most crystal structures (including the opioid receptors structures used here), hydrogen atoms from the docked ligand pose are not considered in the RMSD calculation.

**3e: Rescoring**

Rescoring is a process similar to docking, requiring both a ligand and a receptor. The interactions between the ligand and receptor of the rescored pose are scored using a scoring function. The major difference is that the GA plays a lesser role in rescoring in comparison to docking.

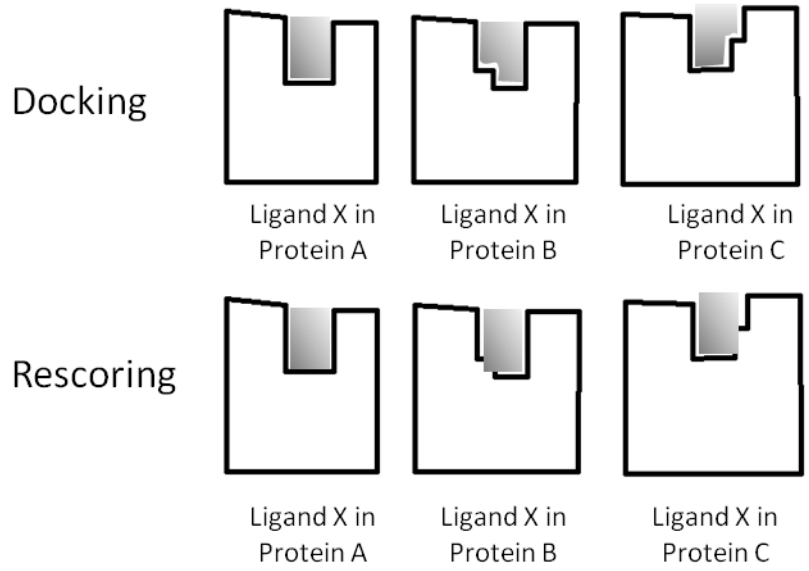



**Figure 22: Schematic comparison between docking and rescoring. The receptor is shown in white with the black outline and the ligands are in shaded grey.**

In docking, the goal of the GA is to search broadly over values of all the rotatable bonds as well as the position and orientation of the ligand in order to optimize the interactions between it and the receptor and minimize steric interactions. However, only small modifications to an existing pose are made during rescoring. These modifications include rotating OH groups to optimize H-bonding, as well as rotating other functional groups on terminal positions to optimize ligand-receptor interactions. Central bonds in the ligand are not rotated during rescoring as this

would cause major conformational changes. Neither does it alter the position or orientation of the ligand (Figure 22).

Rescoring is useful in two scenarios. First, if there is a need to determine the score of a specific pose using a different scoring function than employed for the original docking, then that pose of interest would be rescored using the second scoring function. Second, rescoring is required to obtain the score of a specific pose in a receptor different from that used in the original docking. Rescoring a pose in another receptor is useful for determining how well a specific pose of one receptor can be docked in other subtypes or mutants of the receptor. There is a limitation to rescoring poses in other receptors. Only crystal structure side chain conformations can be used in rescoring in the new receptor. If the original receptor had flexible side chains, then additional discrepancies can exist because the conformation of the side chain is different from that of the crystal structure.

## Materials and Methods

GOLD version 1.6 was used as the docking program for all docking experiments performed. CHEMScore and CHEMPLP were used as the scoring functions[48]. CHEMScore and CHEMPLP were chosen because they have been evaluated as the best scoring functions available in GOLD[68,78]. However, there is no single scoring function that outperforms all the other functions for every ligand-receptor combination[75].

HERMES version 1.6 was used as a visualization interface used to prepare and analyze the docking results. Cygwin was used to run AWK and UNIX scripts used for data processing.

## 1: Crystal Structure Modifications

The crystal structure files of the μ (4DKL)[44], δ (4EJ4)[30] and κ (4DJH)[46] receptor-lysozyme chimeras were downloaded from the RCSB Protein Data Bank (PDB)[41]. All atoms corresponding to ligands were deleted from these files. Furthermore, the lysozyme domain of each chimera was deleted as well. The loop of the receptors where the lysozyme was inserted was reconstructed using ArgusLab[43]. Visual analysis of the receptors showed that some side chains and entire amino acids were missing (Table 7). These were reconstructed using ArgusLab.

**Table 7: Undetected or missing whole amino acids and amino acid side chains in opioid receptors PDB files.**

| μ Receptor: 4DKL | δ Receptor: 4EJ4 | κ Receptor: 4DJH |
|---|---|---|
| Whole Amino Acid | Whole Amino Acid | Whole Amino Acid |
| N-Terminus: | N-Terminus: | N-Terminus: |
| G52,S53,H54,S55,L56,C57,P58,Q59, T60,G61,S62,P63,S64 | G36,S37,P38,G39,A40 | G38,G39,T40,T41,M42,G43,S44,E45, D46,A47,N48,L49,E50,P51,A52,H53,I54 |
| C-Terminus: | C-Terminus: | C-Terminus: |
| P353,T354,S355,S356,T357,I358, E359,Q360,P361 | F329,R330,Q331,L332,C333,R334, T335,P336,C337,G338,R339,Q240, E341,P242 | L348,K349,M350,R351,M352,E353,R354, N355,S356,T357,S372 |
| ICLIII (Lysozyme Attachment Site): | ICLIII (Lysozyme Attachment Site): | ICLIII (Lysozyme Attachment Site): |
| M264,L265,S266,G267,S268,K269 | L245,L246,S247,G248,S249,K250 | S262 |
| ECLIII: | ECLIII: | ECLIII: |
| N/A | N/A | T302,S303,H304,S305,T306 |
| Undetected Sidechains | Undetected Sidechains | Undetected Sidechains |
| M65,T67,K260,R263,R273,N274,R276 | K79,V154,K155,R192,Q201,R241, S242,R244,K252,R291,R292,E323 N324,R327 | T85,F88,K89,K165($NH_2$),L167,F169, K200,D217,R257(guanyl),L259,K265, R267,R270,R271, V296,S301,E335,R342 |

Several commonalities are seen between the three receptors in term of amino acids that are not detected. First, the N and C-terminal sequences are not detected. Second, all three receptors have

a short sequence missing from the ICLIII, due to a deletion made in order to attach the lysozyme domain. The κ receptor has several unique features. First, a five amino acid sequence from the ECLIII was not detected. Second, for one lysine residue (K165), the sidechain butyl group was detected but the amino group of the side chain was not. Likewise, the side chain propyl group, but not the guanyl group, of an arginine (R257) was detected.

The final modification that was executed was the addition of hydrogen atoms using SwissPDB. Since the pKa of protonated histidine is around 6, it was determined that the neutral form of histidine should be used at the physiological pH (*ca* 7.4)[79]. Next, a decision had to be made in choosing the tautomeric form of the imidazole ring to use. The rationale for choosing which tautomer to use is shown below (Table 8). The histidine residues 223, 297 and 319 of the μ receptor, 278 and 301 of the δ receptor, and 291 and 304 of the κ receptor are either in or in close proximity of the main binding pocket.

**Table 8: Tautomeric forms of histidines used and rationale for choosing them.**

| μ Receptor: 4DKL | δ Receptor: 4EJ4 | κ Receptor: 4DJH |
|---|---|---|
| His 54: Hδ Randomly assigned due to rapid tautomerization as a consequence of complete solvent (water) exposure | His 152: Hε H-bond donor to phenolic OH of Tyr 147<br>Nδ exposed to surface | His 53: Hδ Randomly assigned due to rapid tautomerization as a consequence of complete solvent (water) exposure |
| His 171: Hε H-bond donor to phenolic group of Tyr 166<br>Nδ exposed to surface | His 278: Hε No H-bond formation, surrounded by hydrophobic aromatic groups | His 153: Hδ H-bond donor to carbonyl O of Ile 158<br>Nε exposed to surface |
| His 223: Hδ- H-bond donor to side chain OH of Thr 225<br>Nε exposed to surface | His 301: Hδ- H-bond donor to phenolic OH of Tyr 109<br>Nε faces inside the binding pocket | His 291: Hε Randomly assigned as a consequence of no H-bond formation, surrounded by hydrophobic groups |
| His 297: Hε H-bond donor to carbonyl O of Ala 240 | | His 304: Hε H-bond donor to side chain of Glu 50 |
| His 319: Hδ- H-bond donor to phenolic OH of Tyr 128<br>Nε exposed to surface | | |

The N-terminus (pKa=~10) as well as the side chains of lysine (pKa=~10) and arginine (pKa=~12.5) were set to their protonated forms. Likewise, the C-terminus and aspartic and glutamic acid side chains were set to their deprotonated forms (pKa=~4.5).

## 2: Binding Site Parameters

The binding sites of the receptors were marked by choosing an atom in the centre of the binding pocket of the receptor and allowing GOLD to search within a set radius of that atom (Table 9).

**Table 9: Identities of the central atoms of the binding site search radii used for all docking experiments.**

| Receptor | Amino Acid Residue | Atom |
|---|---|---|
| κ: 4DJH | Asp 138 | Carboxylic Acid O (Oδ2) |
| δ: 4EJ4 | Tyr 308 | Phenolic O |
| μ: 4DKL | Asp 147 | Carboxylic Acid O (Oδ2) |

The search radii selected for the docking were 30Å and 50Å.

## 3: GA Parameters

All docking experiments used the same set of GA parameters provided in Table 10 (Refer to Sections: 3a,b).

**Table 10: List of the GA parameters.**

| | |
|---|---|
| Population Size | 100 |
| Selective Pressure | 1.1 |
| Number of Islands | 5 |
| Niche Size | 2 |
| Number of Generations | 100000 |
| % Crossover | 95 |
| % Mutation | 95 |
| % Migration | 10 |

## 4: Flexible Side chains

Flexible side chains were used when DAMGO and β-casomorphin-5 and 7 were docked (Table 11). GOLD will permit a maximum of 10 side chains of the receptor to be flexible. Initially, random side chains within the binding site were chosen to be flexible. Visualization of the docking results with the flexible side chains identified the side chains that made contact with the ligand the most frequently among the docked poses. These were selected to being flexible for the subsequent docking experiments.

**Table 11: List of amino acid side chains in receptor binding site allowed to be flexible.**

| Ligand: DAMGO | | | Ligand: β-Casomorphin-5 | | |
|---|---|---|---|---|---|
| μ Receptor 4DKL | κ Receptor 4DJH | δ Receptor 4EJ4 | μ Receptor 4DKL | κ Receptor 4DJH | δ Receptor 4EJ4 |
| Q124 | Q115 | Q105 | Q124 | Q115 | Q105 |
| D147 | D138 | D128 | N127 | V118 | K108 |
| Y148 | Y139 | Y129 | Y128 | Y119 | Y109 |
| M151 | M142 | M132 | D147 | D138 | D128 |
| T218 | S211 | M199 | Y148 | Y139 | Y129 |
| W293 | W287 | W274 | M151 | M142 | M132 |
| V300 | I294 | V281 | V236 | V230 | V217 |
| W318 | Y312 | L300 | V300 | I294 | V281 |
| I322 | I316 | I304 | I322 | I316 | I304 |
| Y326 | Y320 | Y308 | Y326 | Y320 | Y308 |
| Ligand: β-Casomorphin-7 Trial1 | | | Ligand: β-Casomorphin-7 Trial2 | | |
| μ Receptor 4DKL | κ Receptor 4DJH | δ Receptor 4EJ4 | μ Receptor 4DKL | κ Receptor 4DJH | δ Receptor 4EJ4 |
| Q124 | Q115 | Q105 | Q124 | Q115 | Q105 |
| D147 | L135 | K108 | N127 | V118 | K108 |
| Y148 | D138 | Y109 | Y128 | Y119 | Y109 |
| M151 | Y139 | L125 | D147 | D138 | D128 |
| T218 | M142 | D128 | Y148 | Y139 | Y129 |
| W293 | C210 | Y129 | M151 | M142 | M132 |
| V300 | K227 | M132 | V236 | V230 | V217 |
| W318 | W287 | W274 | V300 | I294 | V281 |
| I322 | Y312 | I304 | I322 | I316 | I304 |
| Y326 | I316 | Y308 | Y326 | Y320 | Y308 |

The set of side chain conformations permitted was taken from the library of side chain conformations provided by GOLD. Each of these side chain conformations was examined visually and if an obvious steric clash is seen, then that side chain conformation was excluded

from the list of permitted conformations during the docking. In addition to the conformations from the library, the original (crystal structure) conformation of the side chain was also added as a permitted conformation.

## Results

### 1: Receptor Superpositioning

Since the sequences of the three opioid receptors are very similar, especially in the TMs, it would be logical to assume that they have very similar three-dimensional structures. To verify this, the modified crystal structures (refer to Materials and Methods: Section 1) of the receptors were superimposed. The criterion for this superposition was that all amino acids from the TMs that are conserved in all three receptors were aligned. If an amino acid was conserved in two of the three receptors, it was not included in the alignment. Likewise, conserved amino acids in the loops were not included as these loops are flexible and thus could have high conformational variability between receptors. 127 amino acid residues were aligned among the three receptors.

**Table 12: RMSD values of the backbone atoms (N, $C_\alpha$, C and O) of each of the receptor pairs. The terminal regions of the sequences were not included in the calculation because of their varying amino acid lengths.**

| Receptor Pair | RMSD (Å) | RMSD (Å) excluding ICLIII |
|---|---|---|
| μ and δ | 3.8 | 1.8 |
| μ and κ | 3.3 | 2.7 |
| δ and κ | 4.6 | 2.9 |

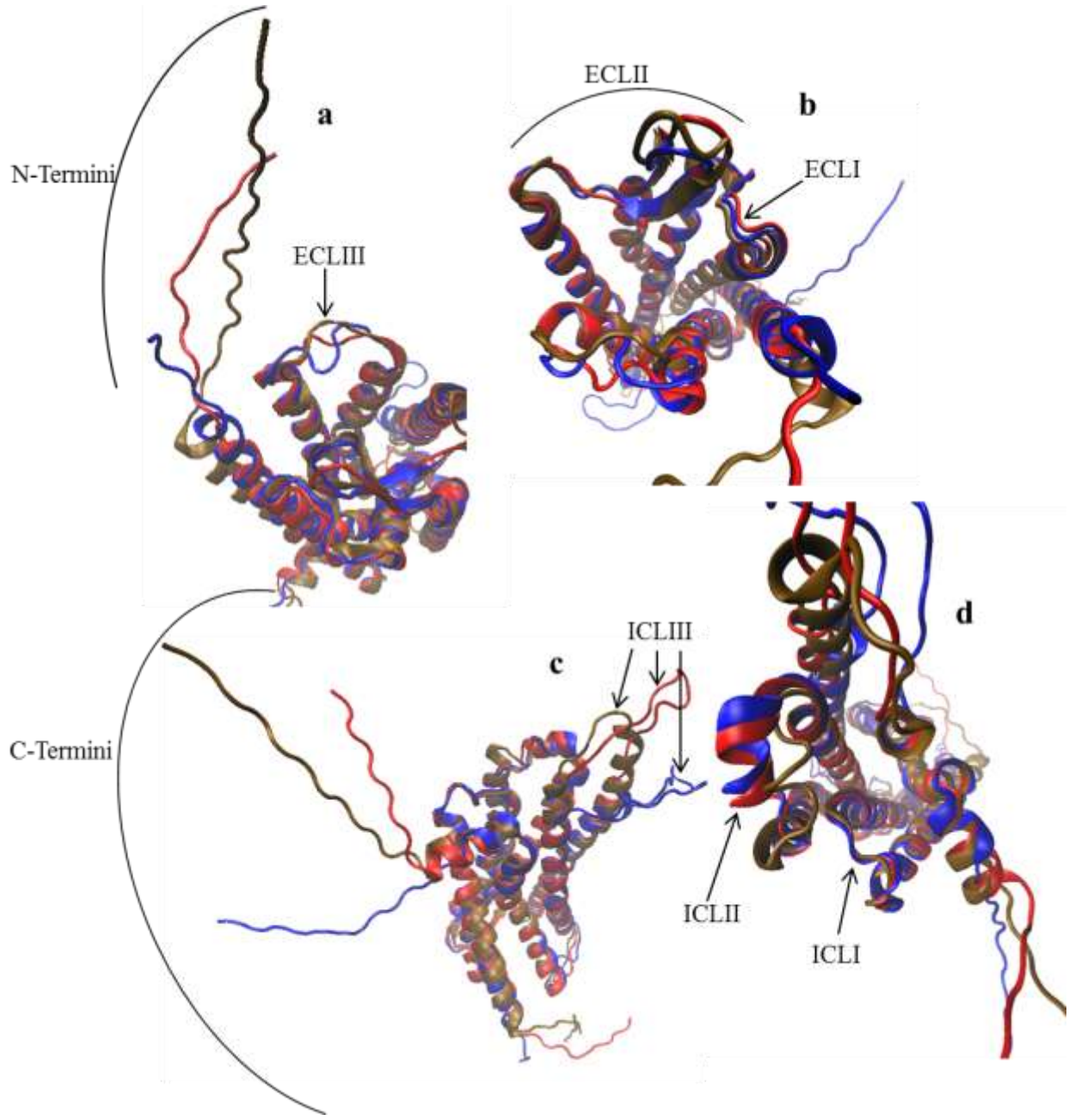



**Figure 23: Superposition of the μ (red), δ (blue) and κ (orchre) receptors: (a) and (b) Viewed from the extracellular face, (c) and (d) viewed from the intracellular face.**

It can be seen (Figure 23a) that the N-terminal region of the receptors is not structurally conserved due to the fact that these N-terminal regions were reconstructed (refer to Materials and Methods: Section 1 for reconstruction details). Further, ECLIIIs are rather dissimilar. It is also important to note that the ECLIII of the κ receptor was reconstructed and thus may not represent

the actual structure of the loop. The ECLs I and II are much more structurally conserved among all three receptors (Figure 23b).

On the intracellular face of the receptors, a similar phenomenon is seen at the C-termini. These segments are structurally dissimilar as they were also reconstructed (Figure 23c). The ICLIIIs are dissimilar, again as a result of reconstruction from the lysozyme attachment site, also evident by its influence of the RMSD values (Table 12). The ICLs I and II are highly conserved structurally (Figure 23d). A small amount of deviation is seen in the ICLII of the κ receptor relative to the μ and δ receptors.

The binding pocket is located among the three ECLs. Since the ECLIII is the most structurally dissimilar of the three ECL among the three receptors, it can be hypothesized that the ECLIII will play a role is the β-casomorphin selectivity. The ICLs are located far away from the binding pocket and across a membrane and thus are unlikely to play any role in ligand selectivity.

## 2: Native Docking

In order to verify that GOLD is an adequate docking program for opioid receptors and whether CHEMPLP or CHEMScore is the most appropriate to use scoring function, native dockings were performed as a positive control, where the crystal structure ligand was docked back into the same receptor.

### 2a: δ Receptor

Naltrindole (Figure 12) was the ligand in the crystal structure of the δ receptor. Naltrindole has two rotatable bonds, indicated by black arrows in Figure 24. Since hydrogen

atoms were not detected, it was unknown whether the tertiary amine of naltrindole was in its neutral or protonated (cationic) form. Initially, the neutral (uncharged) form of naltrindole was used for the docking and docked 100 times into the δ receptor (Figure 24).

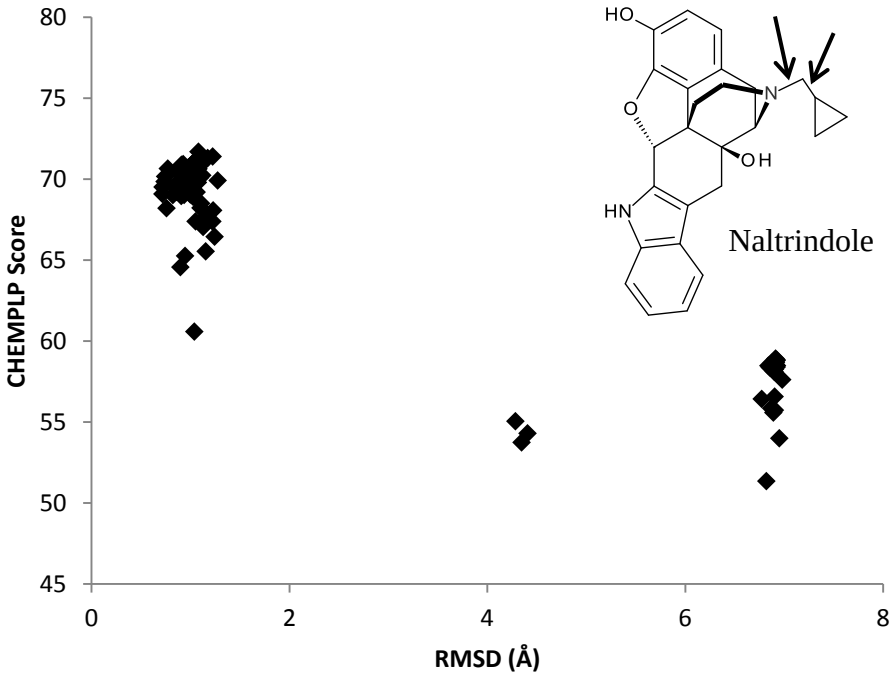



**Figure 24: Comparison of the CHEMPLP score and the RMSD of 100 dockings of neutral naltrindole to the δ receptor using a 30 Å search radius.**

The score gives the predicted quality of the docking result and the RMSD gives the actual quality of the docked pose by comparing its coordinates to that of the crystal structure. Therefore, if GOLD produced good poses, those with low RMSD should have high score. It can be seen that most of the docked poses form a cluster in the region of high CHEMPLP score (65-75) (refer to Introduction: section 3cii) and a low RMSD (*ca* 1.0 Å) (refer to Introduction: section 3d). Two smaller clusters can be seen; one at *ca* 4.5 Å RMSD and another at *ca* 7.0 Å RMSD. An important observation is that there are no "false positives" in the graph. A false

positive occurs when a pose is found that has a high RMSD and high score. This means that GOLD is scoring incorrect poses as high as the correct poses. “False negatives” can also occur where a pose is found with a low RMSD and low score. This means that GOLD is scoring correct poses as low as incorrect poses. Neither false positives nor false negative are seen in Figure 24.

Next, the same docking experiment was repeated expect for naltrindole in its protonated cationic form making a cationic species.

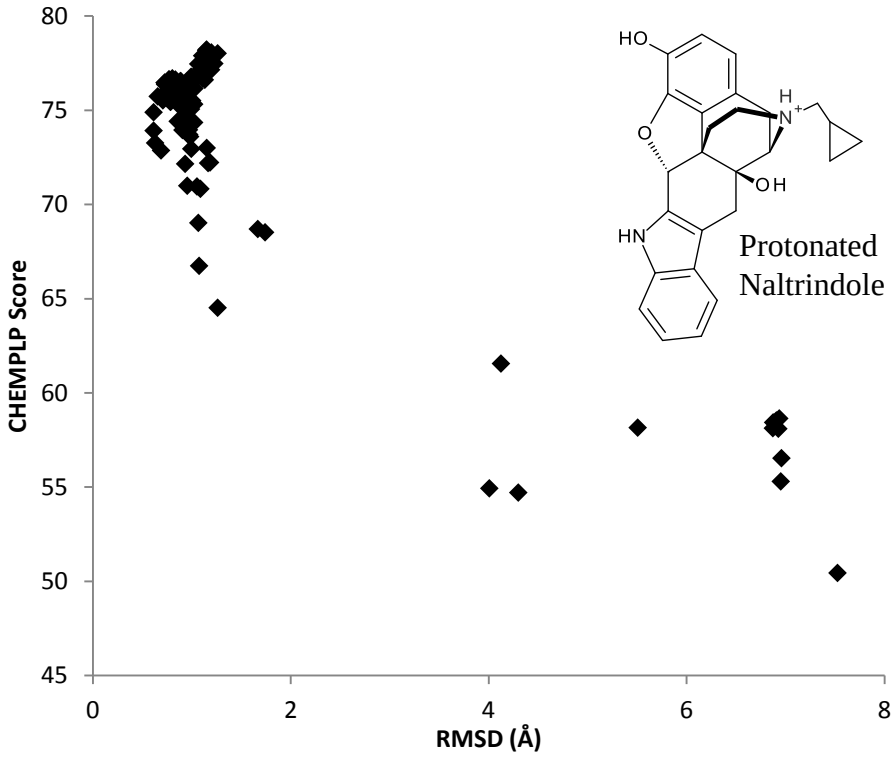



**Figure 25: Comparison of the CHEMPLP score and the RMSD of 100 dockings of protonated naltrindole to the δ receptor using a 30Å search radius.**

The three clusters (*ca* 1.0, 4.5 and 7.0 Å RMSD) were also seen in the docking poses of the protonated naltrindole (Figure 25), however, the RMSD distributions of the relative frequency of occurrence are different for the two native docking experiments (Figure 26).

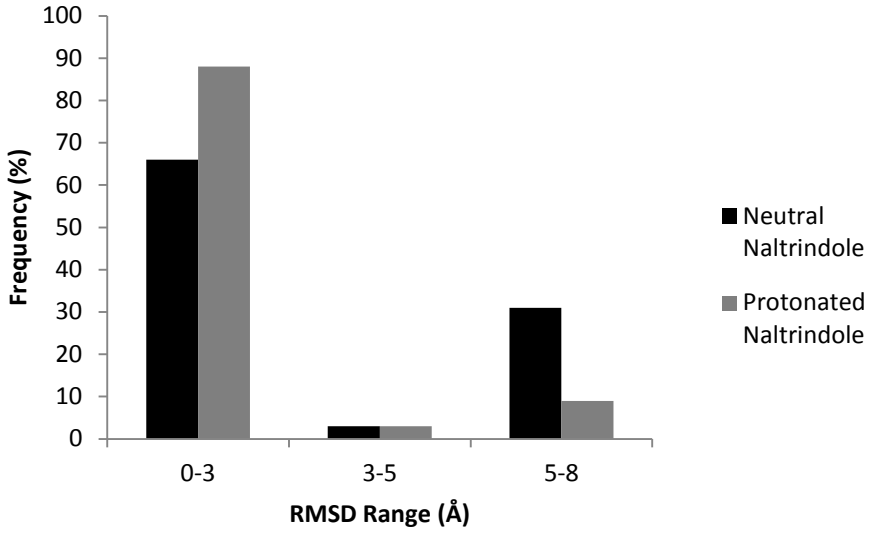



**Figure 26: Frequency of occurrence of 100 naltrindole docked poses within various ranges of RMSD to the crystal structure.**

The protonated naltrindole had a larger population of poses with a low RMSD value compared with the neutral naltrindole. This was a clear indication that the presence of the proton and/or formal charge of the amine promote the correct docking pose of naltrindole.

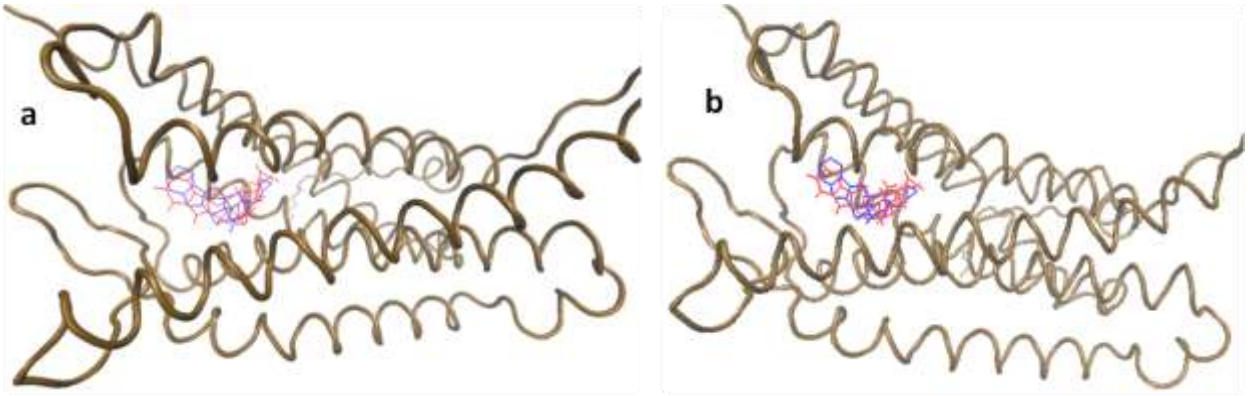



**Figure 27: Highest scoring docked poses (red) of a) neutral naltrindole (score=71.7, RMSD=1.09Å) and b) protonated naltrindole (score=77.9, RMSD=1.10Å) shown with the crystal structure pose (blue). The extra substituents of the docked naltrindole are hydrogen atoms.**

The highest scoring docking pose was taken from each of the docking experiments and their RMSD values and scores were compared (Figure 27). The RMSD values of these two poses were virtually identical, however the protonated naltrindole had a higher score relative to the neutral naltrindole which can be seen in the largest clusters of docked poses at RMSD of *ca* 1.0Å (Figure 24 and Figure 25).

The reason that the protonated naltrindole had a higher score is that the proton interacts with Asp128. While this interaction is considered as a salt bridge, CHEMPLP resisters it as a type of H-bond (refer to Introduction: section 3cii) (Figure 28). The formation of this extra interaction provided the increased score, which also helps to position the naltrindole correctly. It should also be noted that this aspartate residue is conserved in all three receptors (δ, μ and κ).

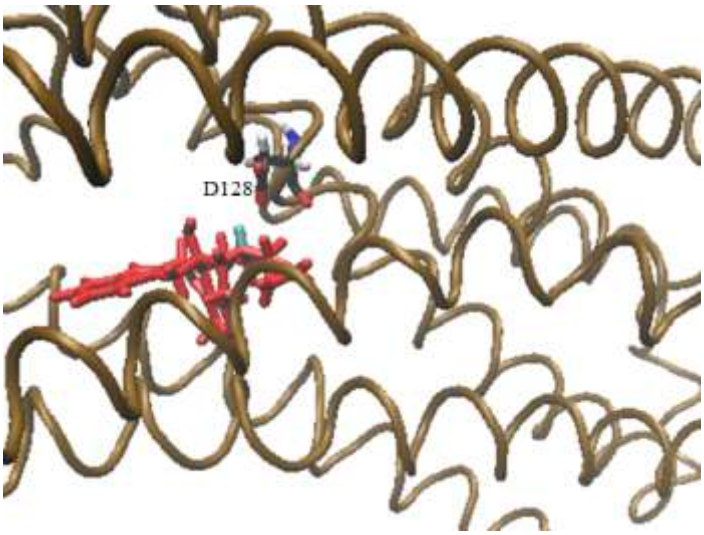



**Figure 28: Salt bridge formation between ammonium centre (proton in cyan) of naltrindole (red) and Asp128.**

Based on these results, all subsequent docking experiments containing basic amines were docked exclusively in their protonated forms.

**2b: μ Receptor**

As mentioned previously (Introduction: Section 1f), the μ receptor ligand, β-FNA (Figure 12), is covalently linked to the side chain of Lys233. GOLD has the ability to introduce a covalent linkage between the receptor and ligand. This option was utilized to accurately represent the ligand-receptor complex. β-FNA has seven rotatable bond indicated by the black arrows in Figure 29. The ligand was docked 100 times into the μ receptor.

The covalent restraint on the ligand clearly improves the RMSD distribution compared with the δ receptor docking using naltrindole (Figure 29): only one cluster is seen (*ca* 1.5 Å RMSD), with only one outlier. It should be noted that naltrindole and β-FNA have the same core structure (Figure 12).

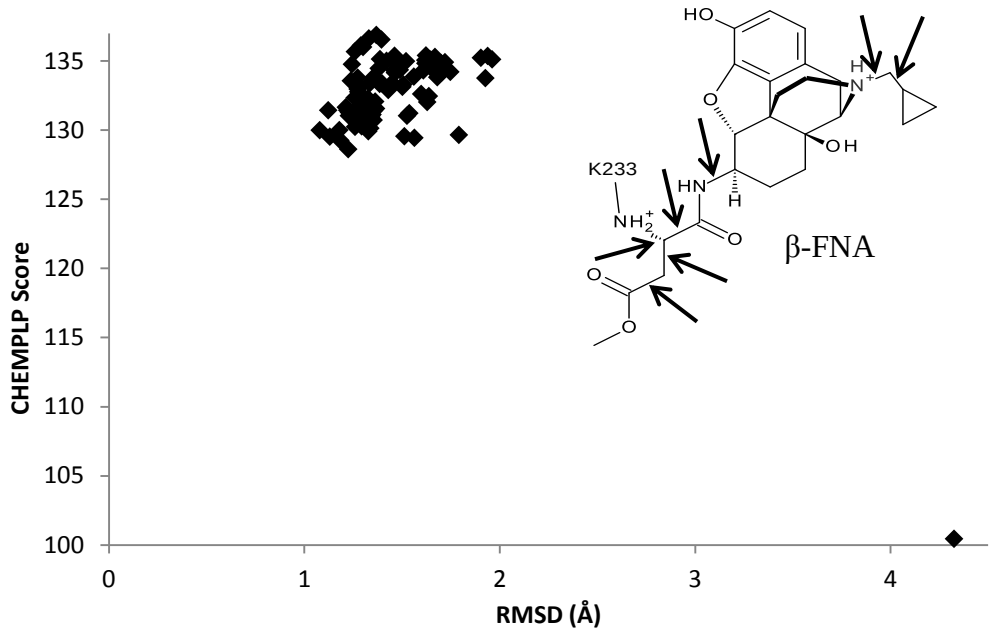



**Figure 29: Comparison of the CHEMPLP score and the RMSD of 100 dockings of protonated β-FNA with the μ receptor using a 30 Å search radius.**

The highest and lowest scoring poses were compared to that of the crystal structure (Figure 30).

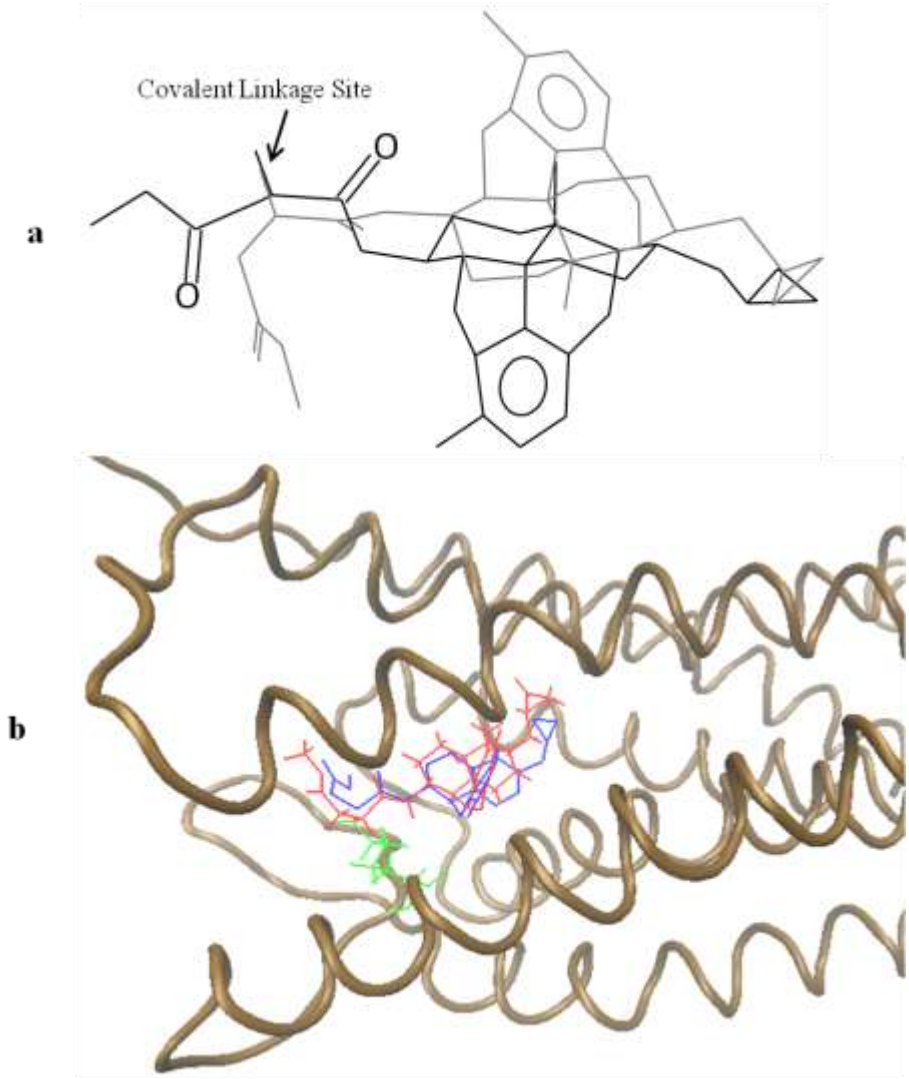



**Figure 30: Lowest and highest scoring docked poses of β-FNA (a) Lowest scoring dock pose (black) (score=100.4, RMSD=4.32Å) shown with the crystal structure (grey) (b) highest scoring docked pose (red) (score=137.2, RMSD=1.36Å) shown with the crystal structure pose (blue) covalently linked to Lys233 (green). The extra substituents of the docked β-FNA are hydrogen atoms.**

The lowest scoring pose has the position of its covalent linkage site the same as the crystal structure as well as the cyclopropyl group. However, the ring system is inverted between the docked pose and crystal structure resulting in a high RMSD (Figure 30a). In the highest scoring pose, the ester group of the β-FNA in the crystal structure makes no contact with the receptor (Figure 30b), consequently, it is nearly impossible to match its conformation with that in

the crystal structure using this docking protocol. This is because GOLD will attempt to positions the ester to make contact with the receptor, which in turn improves the score. Nonetheless, the core structure of this ligand superimposes very well.

Since naltrindole and β-FNA have the same core structure, it would be reasonable to expect that these core structures have very similar poses in their respective receptors.

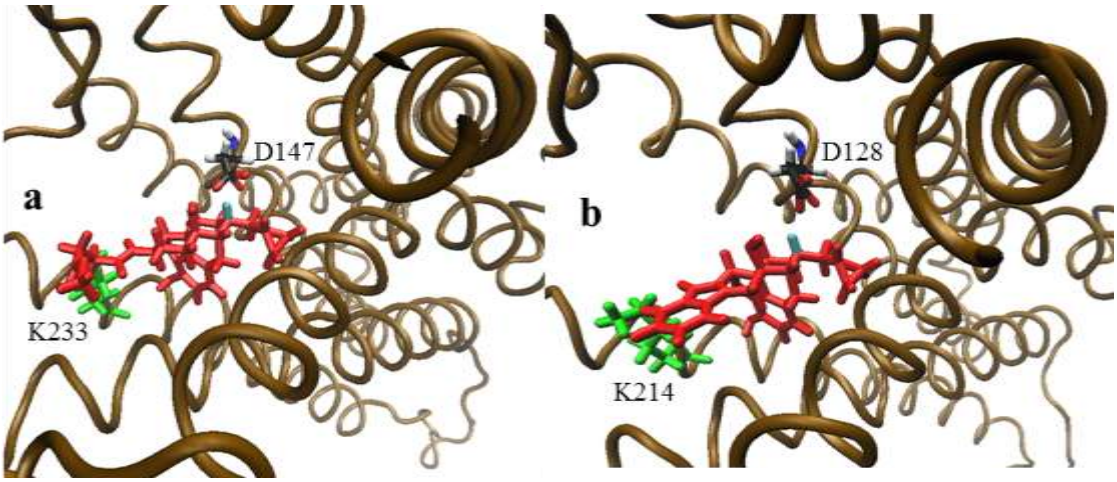



**Figure 31: Comparison of top docking poses of a) β-FNA (red) docking to the μ receptor with covalent linkage to Lys233 (green) and b) naltrindole (red) docking to the δ receptor with Lys214 (green) (ammonium centre proton in cyan).**

Indeed, the core structures of these two alkaloid ligands do orient themselves in a very similar pose (Figure 31). Both protons of the ammonium centres face the conserved aspartate residue. Likewise, Lys233 that makes the covalent linkage is conserved in all three receptors. Residue Lys214 in the δ receptor is located in the same position relative to its ligand core structure.

**2c: κ Receptor**

The κ receptor ligand, JDTic (Figure 12), is considerably different from the ligands of the μ and δ receptor. Whereas the ligands of the μ and δ receptors have very rigid core structures, that of

the κ receptor, a peptidomimetic, is considerably more flexible containing six rotatable bonds indicated by black arrows in Figure 32.

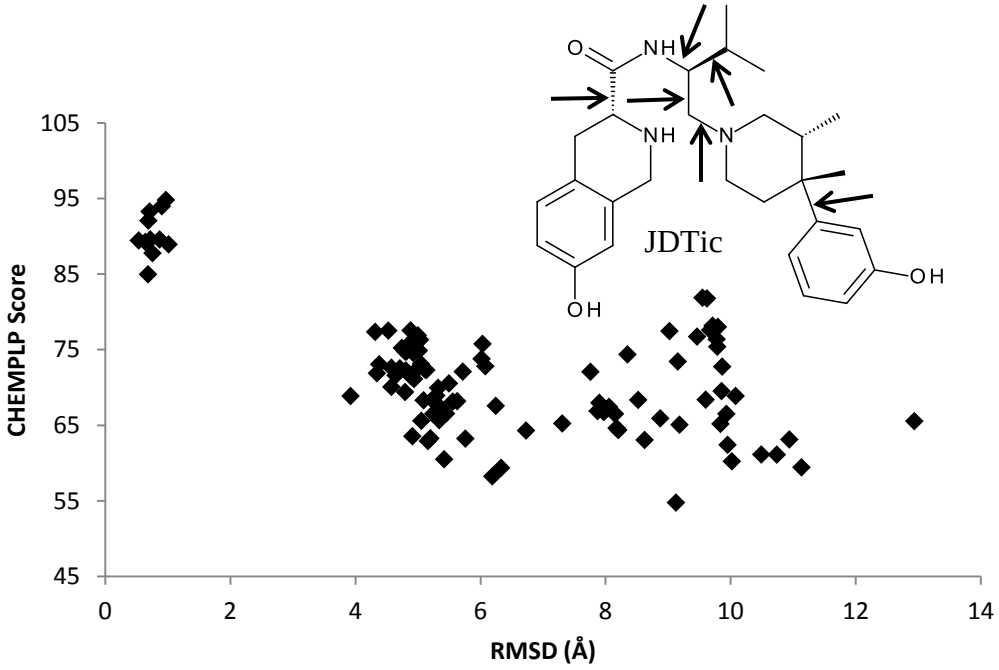



**Figure 32: Comparison of the CHEMPLP score and the RMSD of 100 dockings of protonated JDTic with the κ receptor using a 30 Å search radius.**

The ligand, JDTic, unlike naltrindole and β-FNA, has a much broader distribution of RMSD values (Figure 32). Three clusters can be seen, one at *ca* 1.0 Å, 5.0 Å, and 9.0 Å RMSD, respectively. Unlike what was seen with naltrindole (Figure 24 and Figure 25), the cluster at *ca* 1.0 Å RMSD was the minor cluster. Nonetheless, this cluster consisted of poses consistent with the crystal structure (Figure 33).

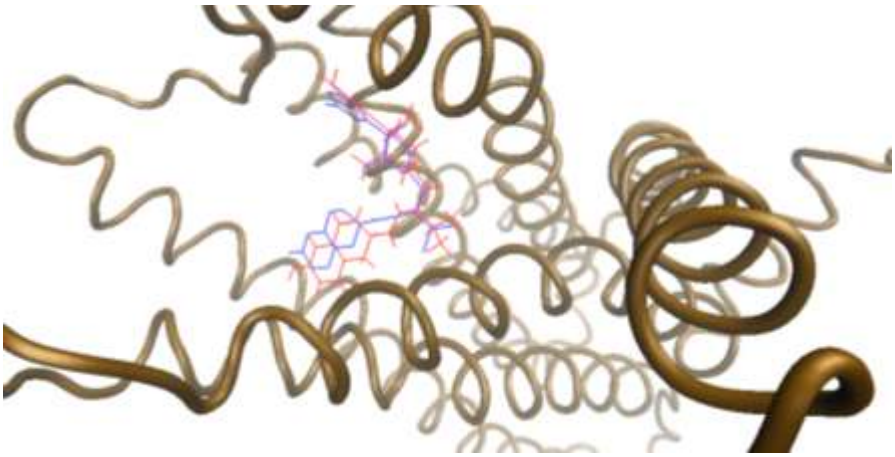


**Figure 33: Highest scoring docked poses (red) of JDTic (score=94.8, RMSD=0.97 Å) shown with the crystal structure pose (blue). The extra substituents of the docked JDTic are hydrogen atoms.**

The poses of the other two clusters (4-12 Å RMSD range) varied quite significantly. This indicates that processing the docking results of the even more flexible β-casomorphin peptide ligands could be very difficult.

**2d: 50Å Radius Docking**

In order to evaluate whether a larger search radius would result in any improvement of the score and/or RMSD, the native ligands were docked to the corresponding receptors using a 50 Å search radius.

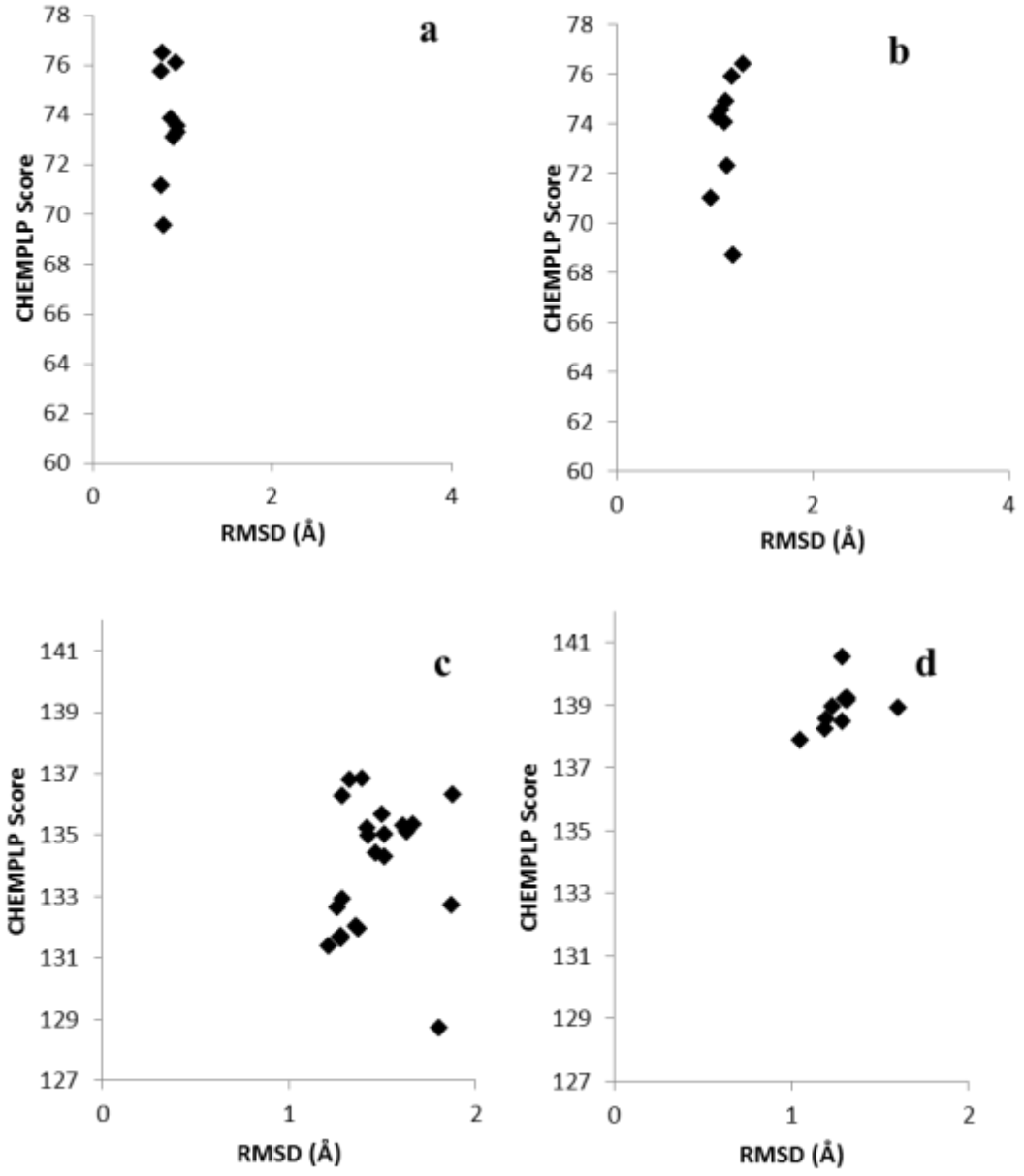



**Figure 34: Comparison of docking results using 30 Å and 50 Å search radii. a) Docking naltrindole to the δ receptor with 30 Å radius, b) docking naltrindole to the δ receptor with 50 Å radius, c) docking β-FNA to the μ receptor with 30 Å radius, d) docking β-FNA to the μ receptor with 50 Å radius, e) docking JDTic to the κ receptor with 30 Å radius, and f) docking JDTic to the κ receptor with 50 Å radius.**

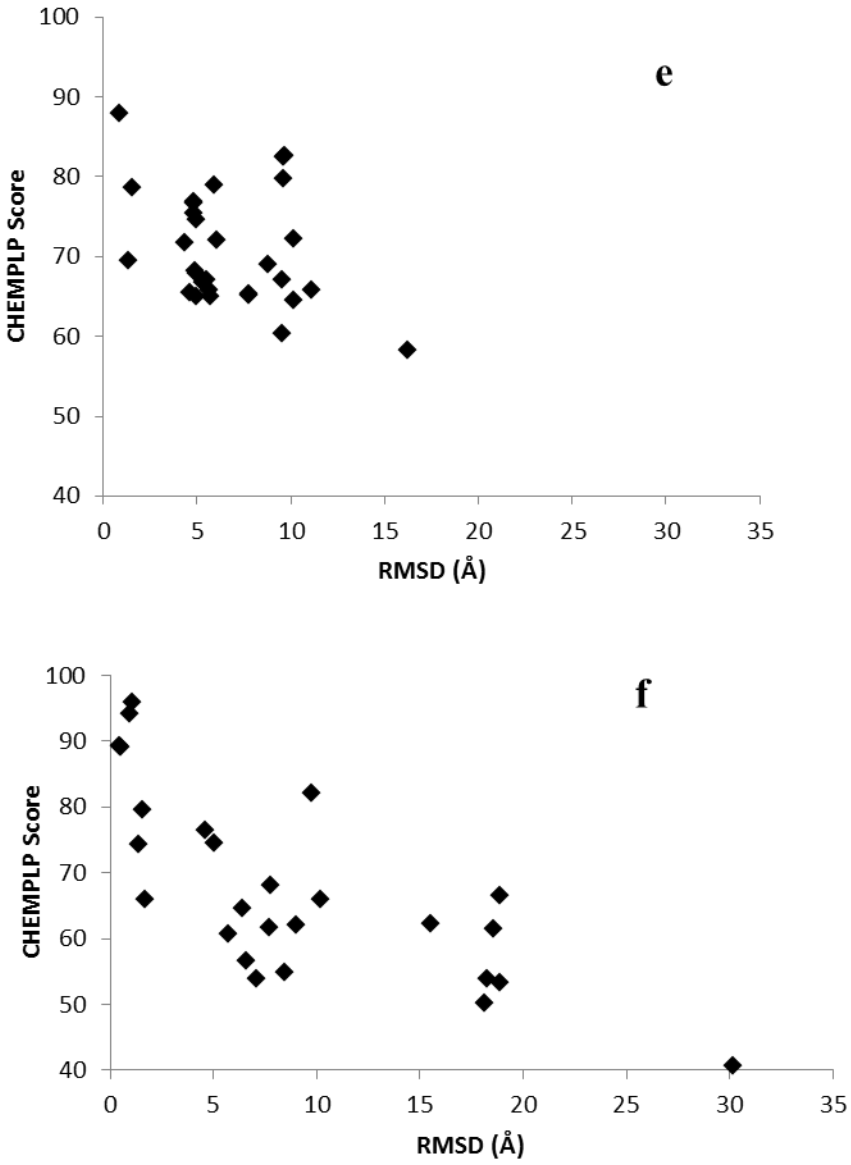



**Figure 32 continued: Comparison of docking results using 30 Å and 50 Å search radii. a) Docking naltrindole to the δ receptor with 30 Å radius, b) docking naltrindole to the δ receptor with 50 Å radius, c) docking β-FNA to the μ receptor with 30 Å radius, d) docking β-FNA to the μ receptor with 50 Å radius, e) docking JDTic to the κ receptor with 30 Å radius, and f) docking JDTic to the κ receptor with 50 Å radius.**

Despite the increased search radius, the RMSD distributions of naltrindole (Figure 34a-b) and β-FNA (Figure 34c-d) are not affected. Nor was there any major change in any of their CHEMPLP score distributions. JDTic, on the other hand, experienced a broader RMSD and

score distribution as a result of the larger radius (Figure 34e-f). Several poses were docked outside the main binding pocket as a result of this excessively large radius (Figure 35).

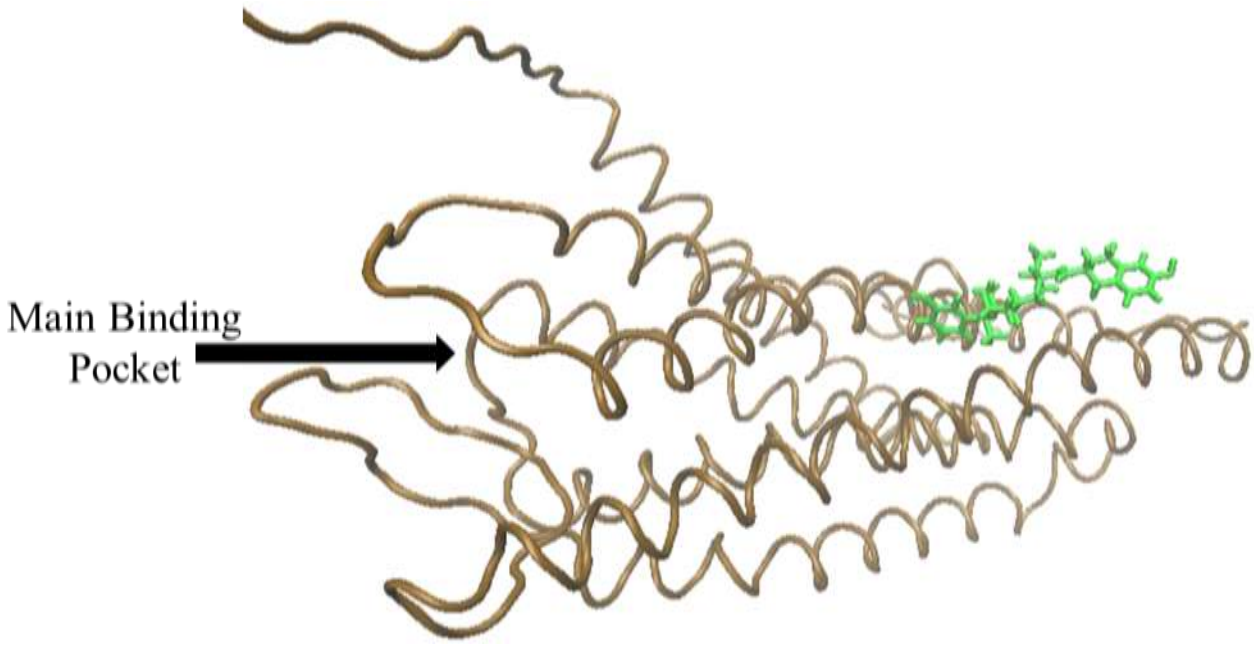



**Figure 35: JDTic (green) docked pose that lies outside the main binding pocket of the κ receptor.**

It was clear from these results that there is no benefit in using a 50 Å over a 30 Å search radius in the docking experiments, and, furthermore, JDTic, which is more flexible than naltrindole and β-FNA, was docked outside the main binding pocket. Docking with the larger search radius, resulted in the computing time per docking to be approximately three times longer as compared to that using the 30 Å search radius. As a result, all subsequent docking experiments used a 30 Å search radius.

Finally, it is important to note that no false positives (refer to Section 2a) were observed in all the docking experiments done using CHEMPLP. This means that for the native docking experiments, GOLD with CHEMPLP is an adequate combination of a docking program and scoring function.

The native docking experiments were repeated using a 30 Å search radius and CHEMScore as the scoring function. What was observed was that CHEMScore was able to discriminate the correct versus incorrect poses through the score of naltrindole in the δ receptor and β-FNA in the μ receptor. However, CHEMScore was not able to discriminate the correct versus incorrect poses of the κ receptor-JDTic complex through the score. Incorrect poses that score equally to that of correct poses are considered false positives. Since JDTic is the closest representation among the native ligand to β-casomorphins due to its flexibility, CHEMScore was not used for any other docking experiments.

## 3: Non-Native Cross Docking

Each of the ligands in the crystal structures, naltrindole, β-FNA and JDTic, is a selective antagonist of their respective receptor. This implies that if each of the crystal structure ligands is docked into the other two receptors (non-native docking), the score of those dockings should be lower relative to that of the native docking. Since these non-native dockings will have no crystal structure as a reference, the only measurement of the quality of the docking is the CHEMPLP score. Performing non-native docking experiments functions as a negative control experiment. That is, can GOLD discriminate between known binders and non-binders to the opioid receptors.

### 3a: Naltrindole

Naltrindole has the highest binding affinity for the δ receptor (Table 2). Therefore, docking of naltrindole to the δ receptor should have the highest score among δ, κ and μ receptors. This is indeed found for the non-native cross docking results (Figure 36). The δ receptor shows a high overall distribution of the CHEMPLP score compared with that of the κ and μ receptors.

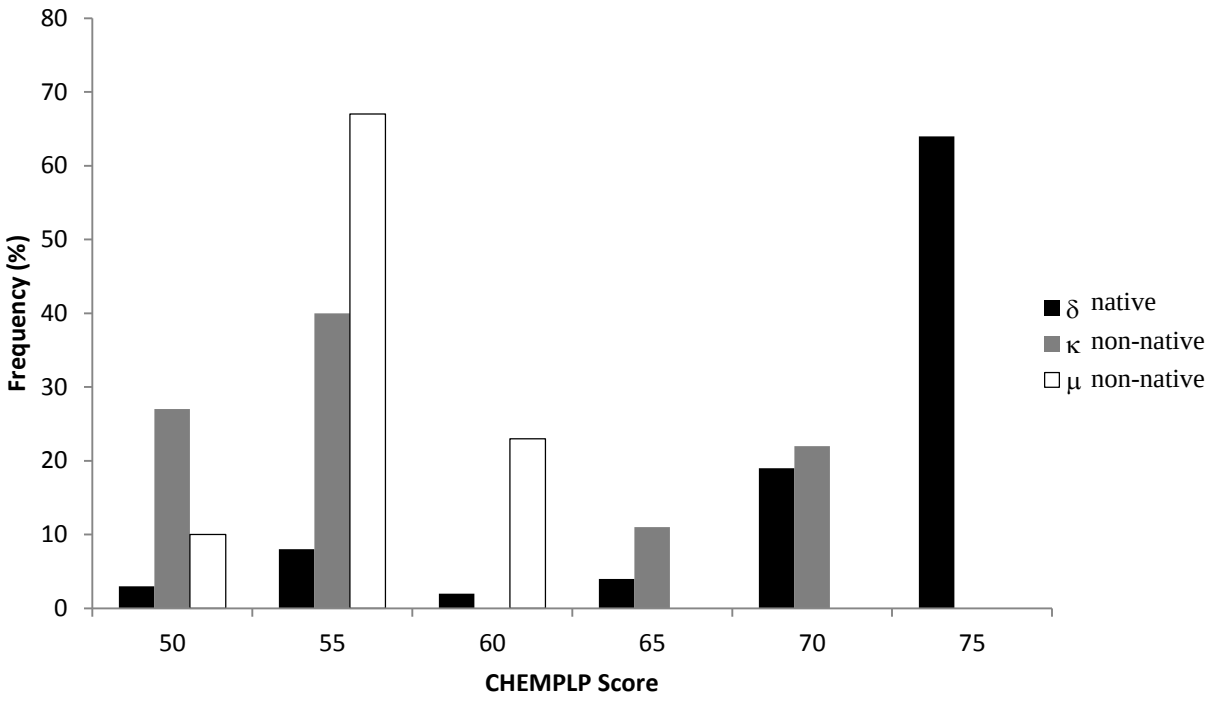



**Figure 36: Comparison of the CHEMPLP score distribution of 100 dockings of naltrindole to the δ, κ and μ receptors.**

**3b: β-FNA**

β-FNA had been shown to bind irreversibly only to the μ receptor, even though the lysine which forms the covalent linkage is conserved in all three receptors[80]. For this docking experiment, the covalent linkage was made for all three receptors. This should mean that the δ and κ receptor dockings should score lower relative to the μ receptor.

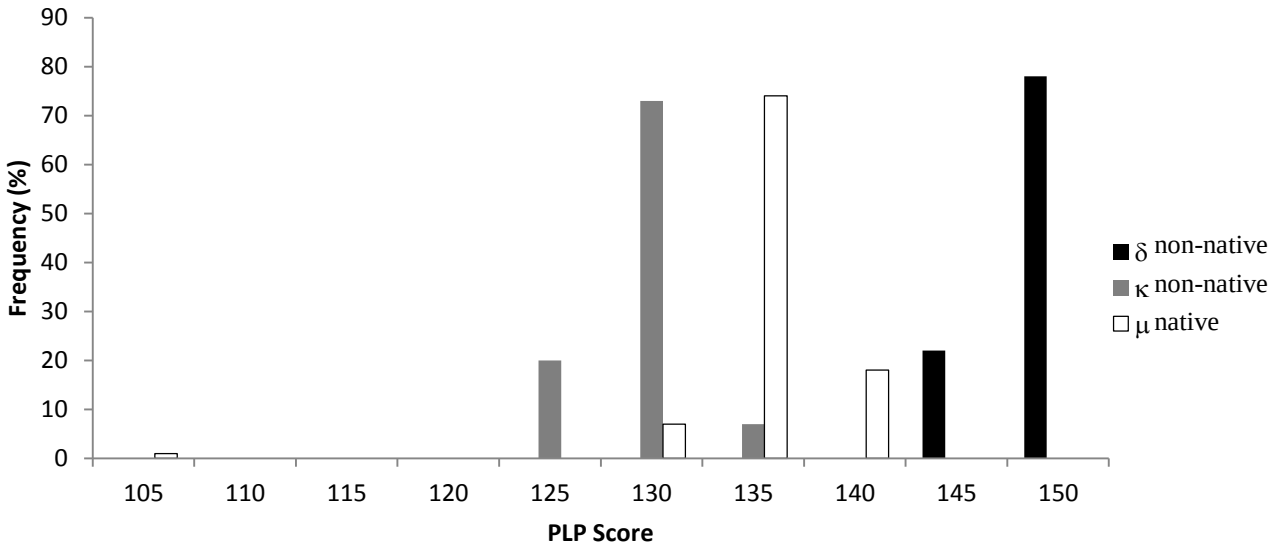



**Figure 37: Comparison of the CHEMPLP score distribution of 100 dockings of β-FNA to the δ, κ and μ receptors.**

However, it was found that while the κ receptor does score lower in comparison to the μ receptor, the δ receptor actually scores higher than the μ receptor (Figure 37). Furthermore, the pose of β-FNA in the δ receptor was nearly identical to that of μ receptor. Therefore, GOLD can discriminate β-FNA between the κ and μ receptors, but not the δ and μ receptors.

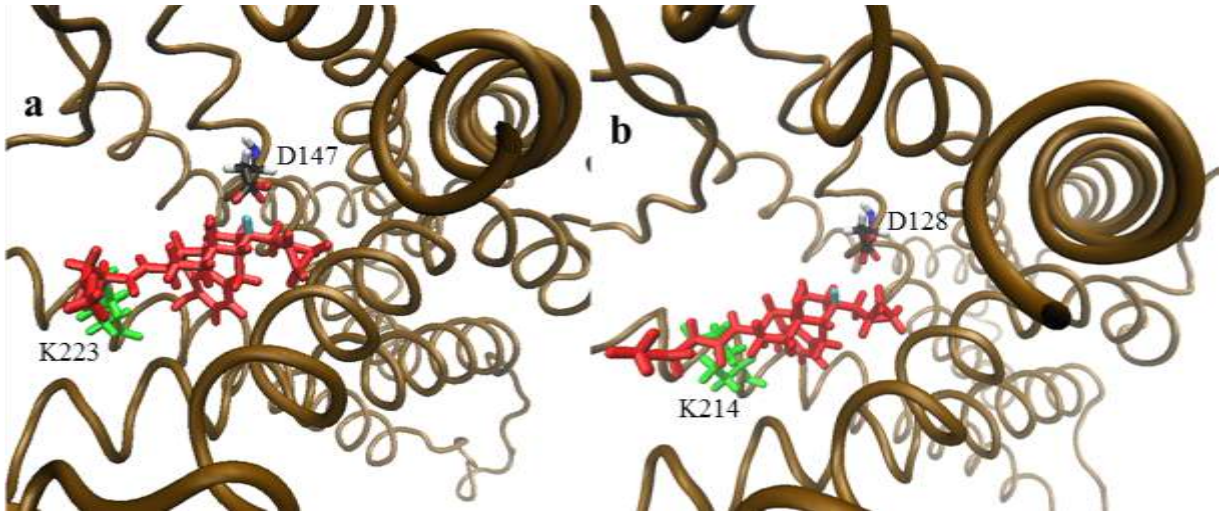



**Figure 38: Highest scoring pose of β-FNA (red) in a) μ receptor and b) δ receptor (ammonium centre proton in cyan). Both receptors have the ligand covalently linked to a lysine (green).**

**3c: JDTic**

Since the flexibility of JDTic led to a broad distribution of the score and RMSD values in the native docking experiments (Figure 32), similar results were expected with the non-native dockings of this ligand. Indeed, similar results were seen when JDTic is docked to the δ and μ receptors (Figure 39).

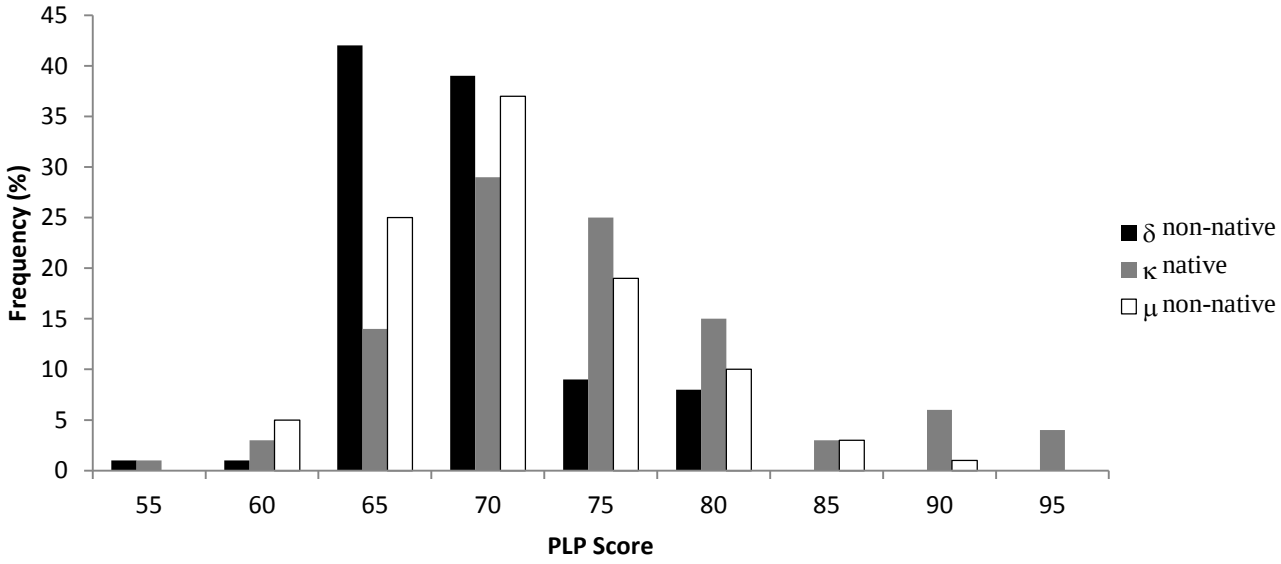



**Figure 39: Comparison of the CHEMPLP score distribution of 100 dockings of JDTic to the δ, κ and μ receptors.**

The δ receptor clearly had a lower average score compared to that of the κ and μ receptors. While the average scores of the κ and μ receptors were very similar, only the κ receptor had docked poses with very high scores (*ca* 95). It can be concluded from these experiments that GOLD with the CHEMPLP scoring function can discriminate between different receptors for very flexible ligands. This provided enough confidence to dock the more flexible β-casomorphins to the opioid receptors and obtain qualitatively meaningful results. As only the

highest scoring poses of JDTic allowed for discrimination among the three receptors, the highest scoring poses of the β-casomorphins should be the focus of analysis.

## 4: β-Casomorphin Docking

The work done by Brantl and co-workers[57] showed that the β-casomorphins bind preferentially to the μ receptor. The group also determined that β-casomorphin-5 has the highest binding affinity among the different β-casomorphins (Table 4).

The three-dimensional coordinates for the β-casomorphin-5 and 7 ligands were generated by ArgusLab.

### 4a: β-Casomorphin-7

β-casomorphin-7 (Table 4), containing 16 rotatable bonds and three tertiary amides, is much more flexible compared to any of the native ligands examined in Section 3. Tertiary amides can undergo 180° flips in the docking algorithm but secondary amides will not. Therefore, the number of dockings per receptor was increased to 500 in order to adequately cover the conformational space. The receptor side chains were initially kept rigid in their crystal structure conformations.

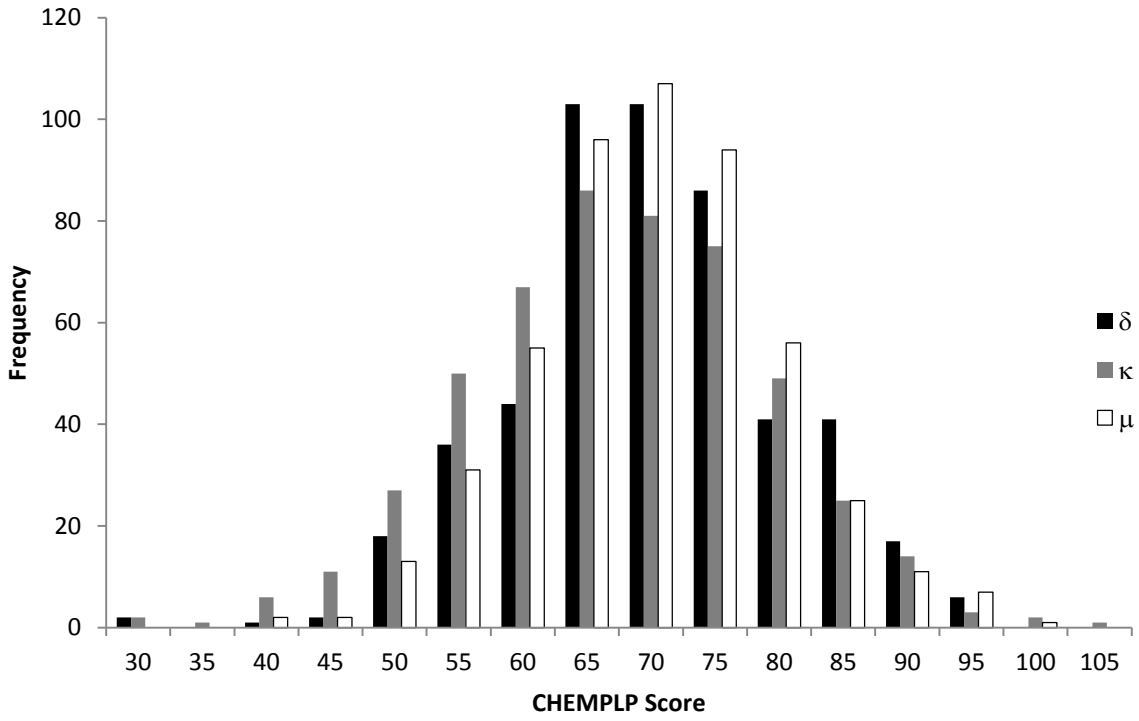



**Figure 40: 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with rigid receptor side chains.**

Despite the success in docking the native ligands to their corresponding receptor, and clear discrimination between receptors in the non-native cross docking experiment with the only exception being the β-FNA docking between the δ and μ receptors, no obvious discrimination is seen between the three receptors in β-casomorphin-7 docking (Figure 40). It is possible that the side chains of the receptor are not positioned in conformations to optimally interact with β-casomorphin-7. Since GOLD has the ability to process flexibility for up to ten side chains of the receptor, flexibility in the receptor was considered in the subsequent docking experiment. First, amino acid residues within the binding pocket of the receptor were randomly selected be flexible (Trial 1) (refer to Materials and Methods: Section 4, Table 11).

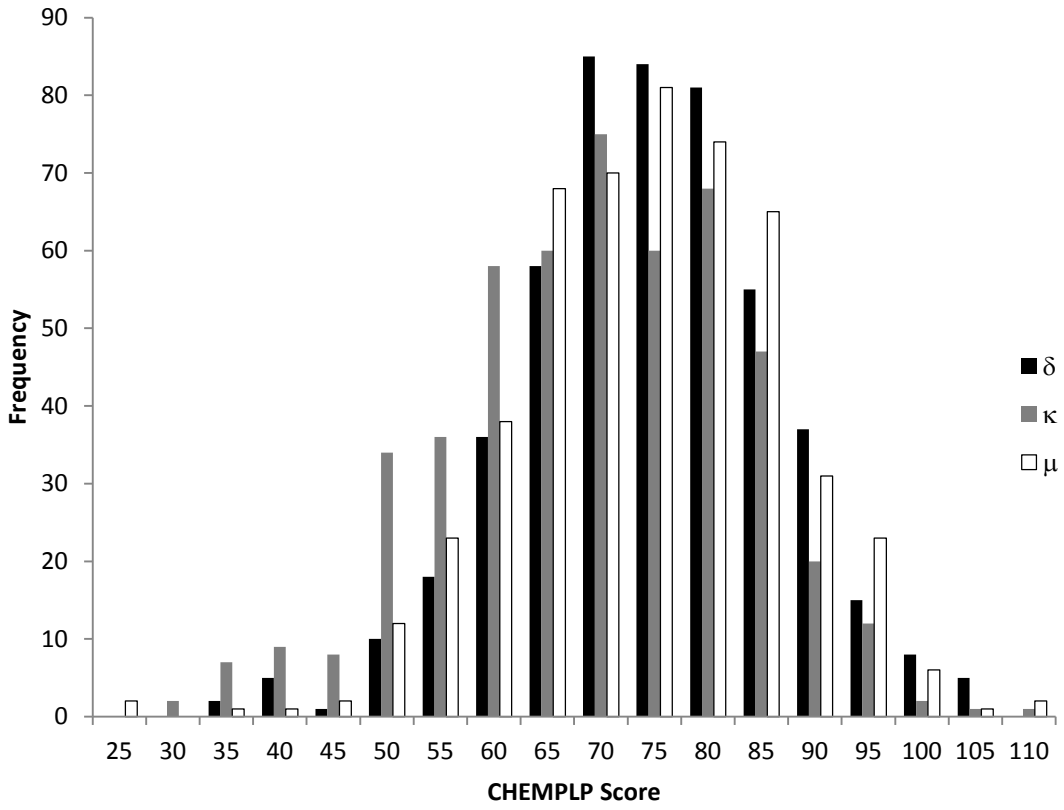



**Figure 41: Trial 1, 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with flexible side chains.**

Though there was some evidence of selection for the μ receptor at the score of 110, these were only two poses for the μ receptor out of the 500. These two poses were not easily seen when graphing the cumulative frequency over the CHEMPLP score (Figure 42).

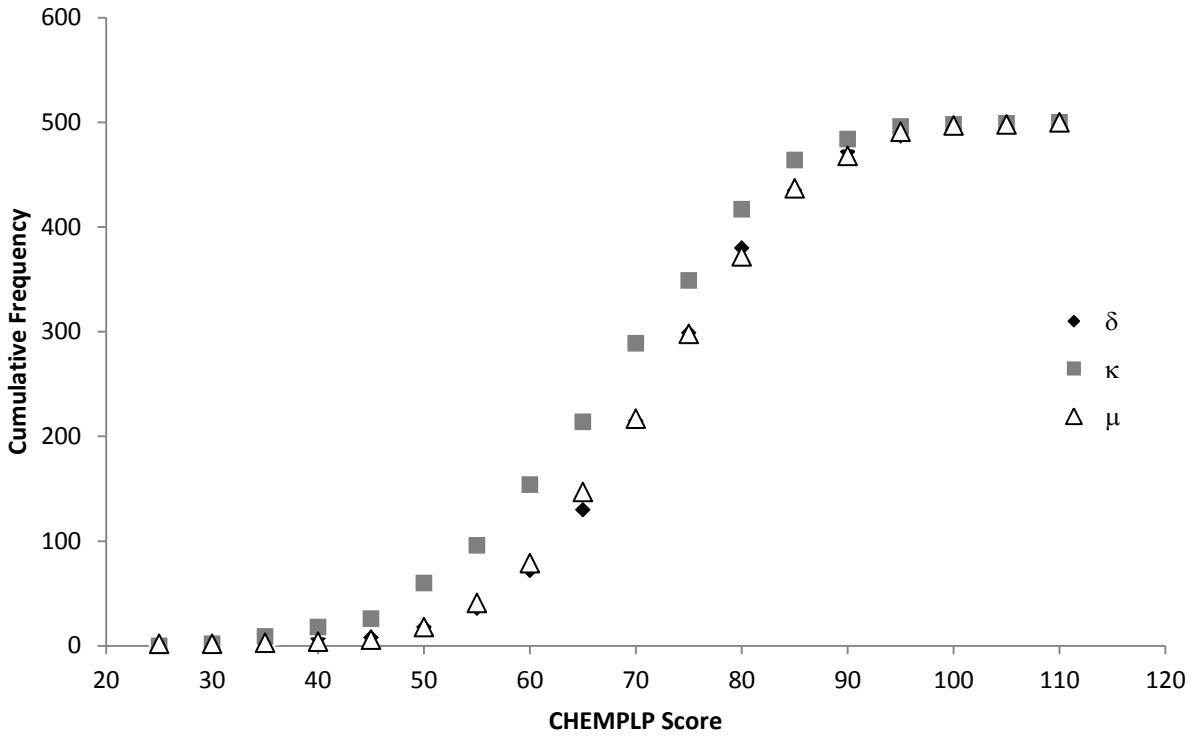



**Figure 42: Trial 1, 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with flexible side chains using cumulative frequency.**

It was concluded that despite these flexible side chains, no clear discrimination between the receptors was observed (Figure 41). However, visual analysis of the docking results showed some amino acid residues of the receptors making contact with the ligand in a wide range of docked poses. It was also found that while tryptophan residues were treated as flexible in trial 1, they did not frequently undergo conformational changes, maintaining its crystal structure conformation, thus they were set to be rigid for all remaining β-casomorphin dockings.

Following from Trial 1 docking results, in Trial 2 (refer to Materials and Methods: section 4, Table 11) the same amino acids in the sequence alignment (Figure 7) were set to be flexible in each receptor. This was held true for all remaining docking experiments involving

flexible side chains. Again, no discrimination among the three receptors was seen in Trial 2 docking (Figure 43). The cumulative frequency graph did not show any difference among the three receptors.

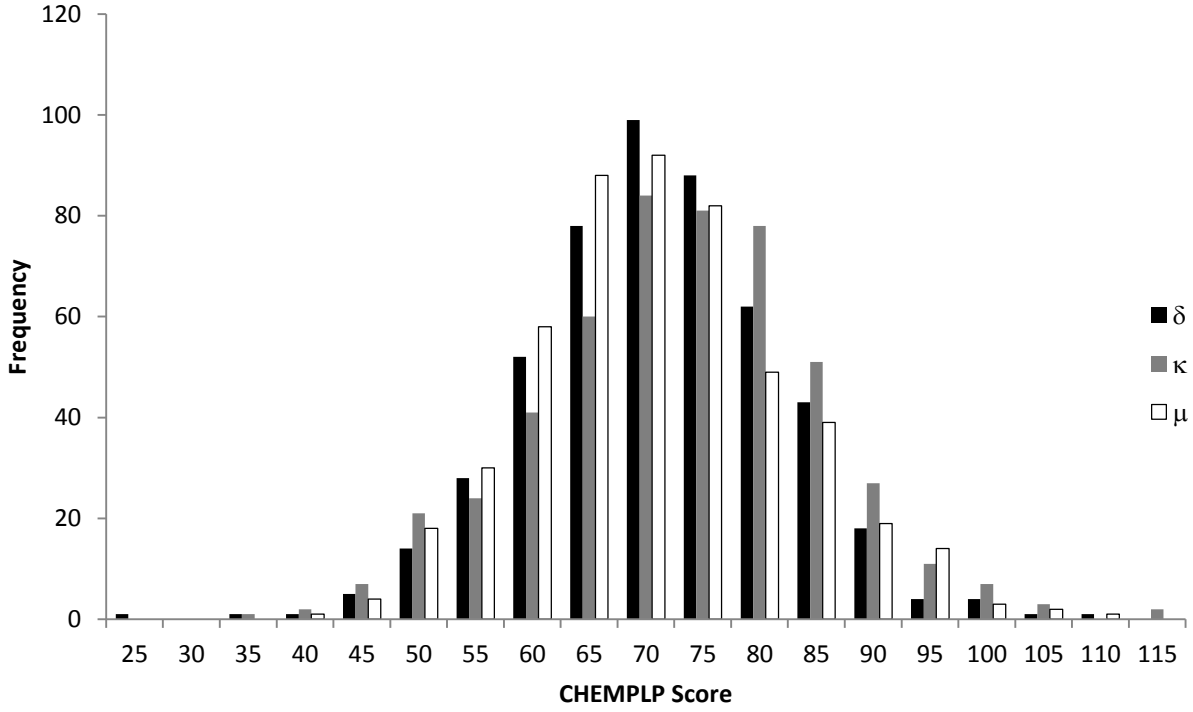



**Figure 43: 500 dockings of β-casomorphin-7 of the δ, κ and μ receptors with trial 2 flexible side chains.**

**4b: β-Casomorphin-5**

β-casomorphin-5 (Table 4) was subsequently used as ligand for docking. The shorter sequence of β-casomorphin-5 means fewer rotatable bonds (11 rotatable bonds and two tertiary amides) and thus fewer possible conformations. β-casomorphin-5 was docked 500 times in each receptor as what was done for β-casomorphin-7 (refer to Section 4a). First β-casomorphin-5 was docked with rigid side chains (Figure 44), which led to no discernible discrimination among receptors in the docking results.

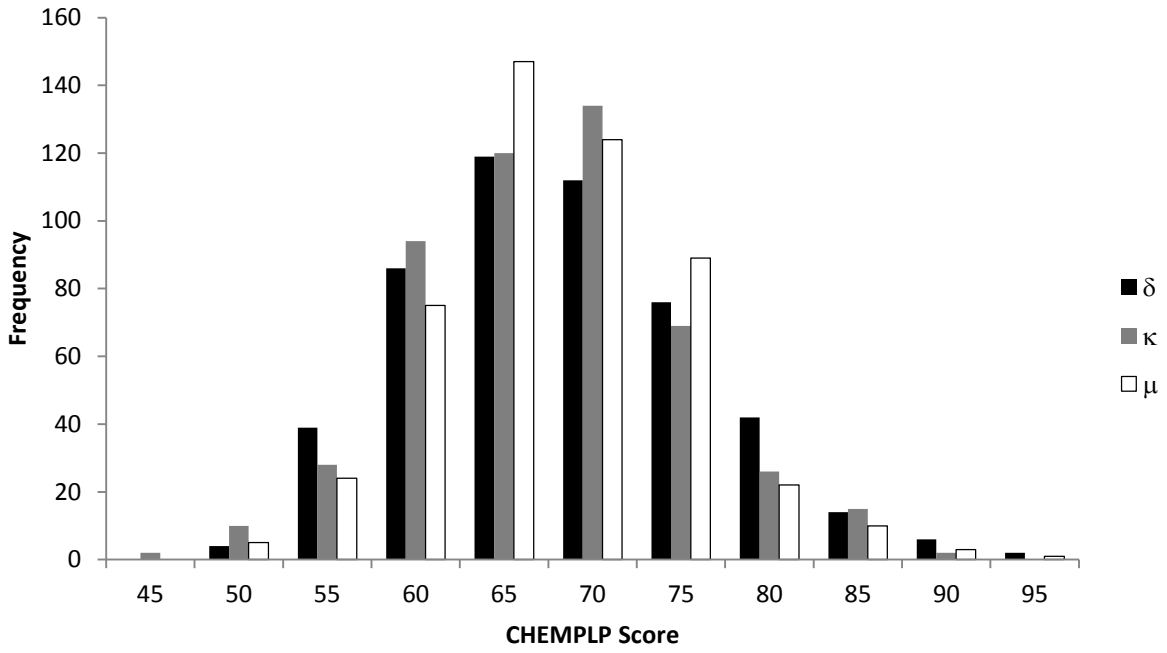



**Figure 44: 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors.**

Next, the docking was carried out again with the same set of flexible side chains as Trial 2 of β-casomorphin-7 (refer to Materials and Methods: section 4, Table 11). The δ and κ receptors were scoring high (70-90 CHEMPLP score) more frequently than the μ receptor (Figure 45). This observation indicates that β-casomorphin-5 appears to favour the δ and κ receptor over the μ receptor under these conditions. This is contradictory to what is shown experimentally based on the work done by Brantl and co-workers[57]; however, this difference is only marginal and not discriminatory as seen in the docking with the native ligands (Figure 36, Figure 37 and Figure 39).

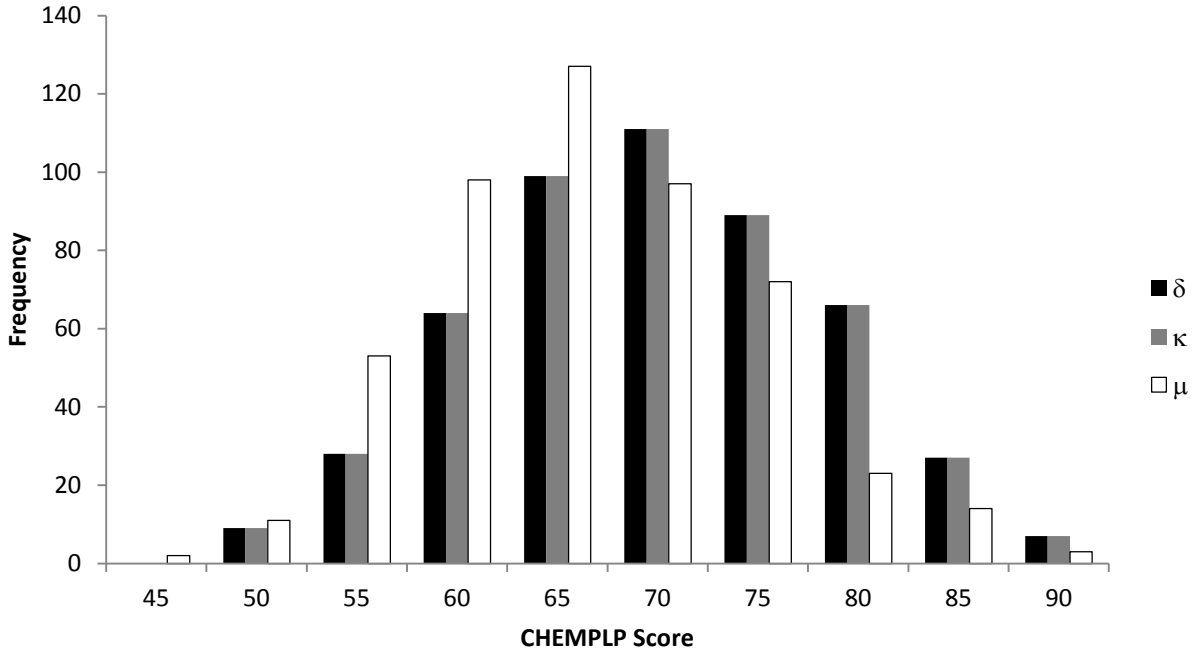



**Figure 45: 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors with flexible side chains.**

A previous molecular dynamics (MD) simulation study done by Borics and Tóth used 2-aminocyclopentane-carboxylic acids (ACPC) in place of the first proline of β-casomorphin of these opioid peptides (Figure 46)[81]. The tertiary amide proline forms can undergo flips; however, the ACPC has locked geometry around the five-membered ring since it will form a secondary amides which does not undergo flips.

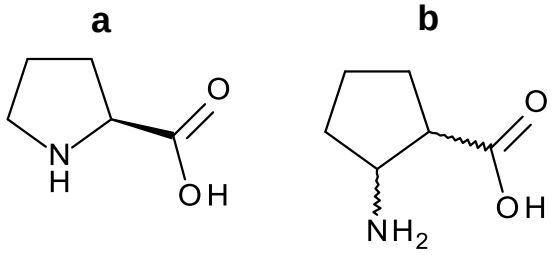



**Figure 46 : Structure of a) L-proline and b) 2-aminocyclopentane-carboxylic acid.**

It was determined that the *cis* configuration of ACPC leads to formation of more stable conformations of the peptides. The authors predicted that when the β-casomorphins bind to the opioid receptors, the tyrosine-proline peptide bond assumes the *cis* conformation[81].

Following this information, the sets of 500 dockings of β-casomorphin-5 with flexible side chains of the receptor were divided based on the conformation of the tyrosine-proline peptide bond (Table 13, Figure 47).

**Table 13: Ratio of *cis* and *trans* Tyr-Pro peptide bond among the δ, κ and μ receptors.**

| | μ Receptor | δ Receptor | κ Receptor |
|---|---|---|---|
| *cis* | 218 | 185 | 204 |
| *trans* | 282 | 315 | 296 |

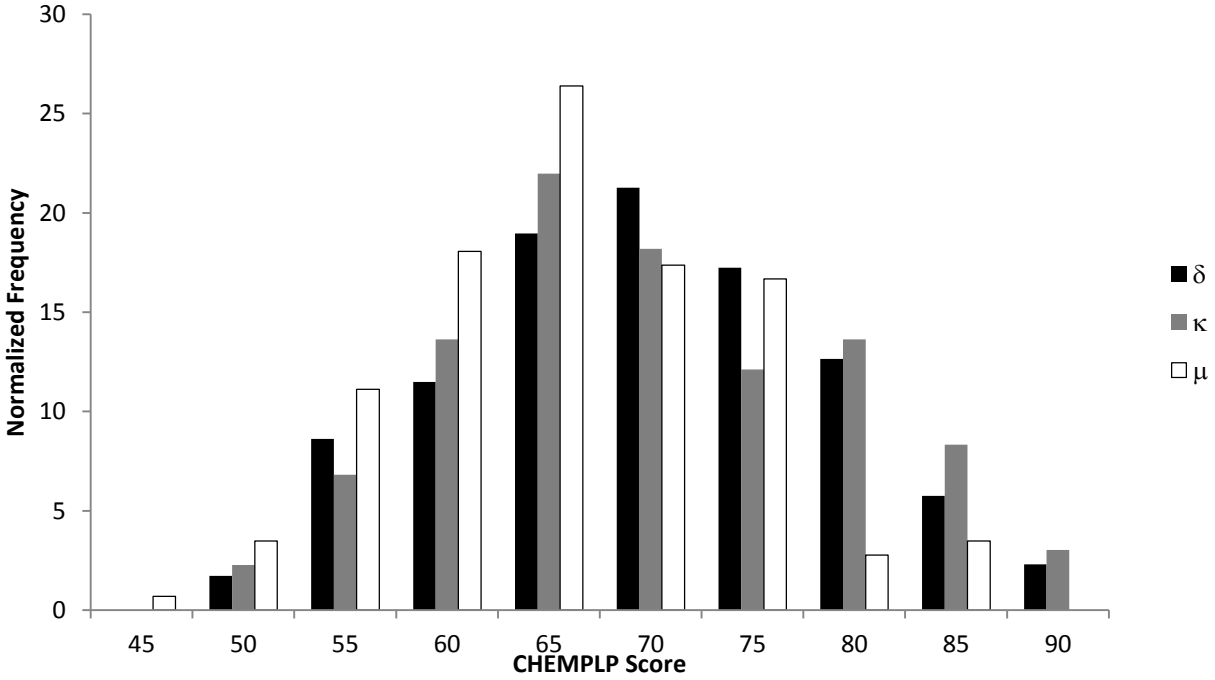



**Figure 47: Subset of the 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors with flexible side chains bearing a *cis* Tyr-Pro peptide bond.**

The subset of dockings bearing the *cis* tyrosine-proline peptide bond were normalized since the ratio of *cis*:*trans* was not the same for each receptor. However, this segregation did not lead to any noticeable difference in the distribution of the CHEMPLP score compared with the entire set of 500 dockings (Figure 45). The *trans* tyrosine-proline peptide bond subset also did not differ from the entire docking set (Figure 48). That is, GOLD docking scores of β-casomorphin-5 were not higher for *cis* versus *trans* conformations of the first Tyr-Pro peptide bond.

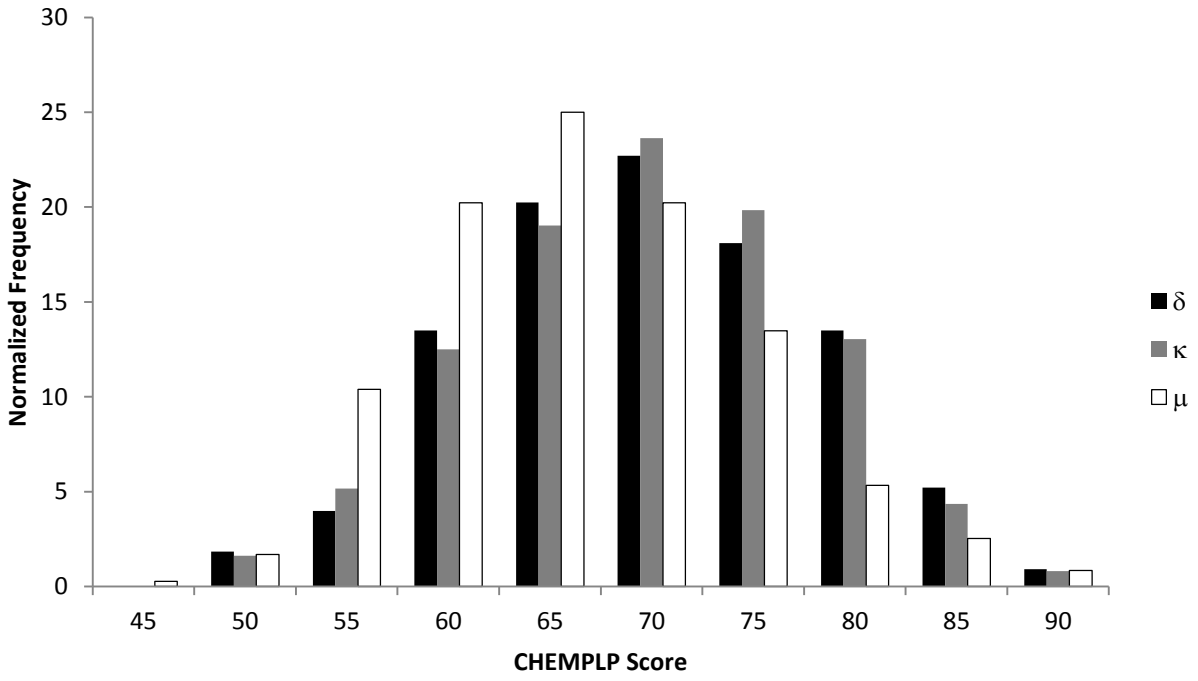



**Figure 48: Subset of the 500 dockings of β-casomorphin-5 of the δ, κ and μ receptors with flexible side chains bearing a *trans* Tyr-Pro peptide bond.**

It was concluded from these β-casomorphin docking experiments that the methodology that was used for the native ligand dockings could not be used to discriminate between opioid receptors for these highly flexible ligands. A different methodology had to be developed to analyze the outcome of the docking results.

## 5: DAMGO Docking

DAMGO (Figure 6) was chosen as a peptide ligand for docking to μ, δ and κ opioid receptors in order to develop a new methodology for analysing the β-casomorphin docking results. Like the β-casomorphins, DAMGO is a peptide, and therefore is highly flexible containing 13 rotatable bonds and one tertiary amide. Further, unlike the β-casomorphins, amino acids of the opioid receptors that are important for DAMGO binding have been studied extensively (refer to Introduction: Section 1gi), although no crystal structure of DAMGO bound to the μ receptor is available. The three-dimensional coordinates for DAMGO were generated using ArgusLab.

### 5a: Traditional Scoring Analysis

The same docking procedures that were used for the β-casomorphins were applied with DAMGO as the docked ligand. First the receptor side chains were kept rigid and as expected, no discrimination was seen between the three receptors (Figure 49), although it is known that DAMGO binds preferentially to the μ receptor.

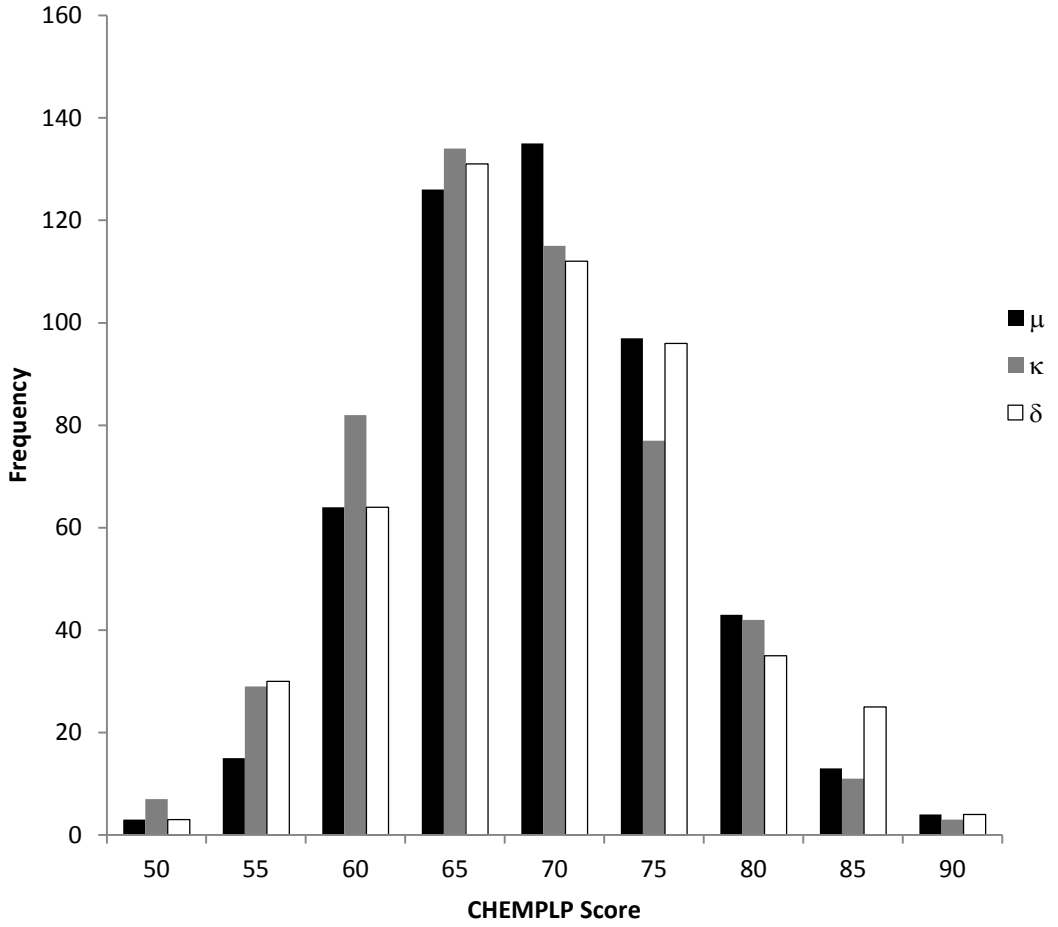



**Figure 49: 500 dockings of DAMGO of the rigid δ, κ and μ receptors.**

Flexibility was then introduced to the side chains of the receptor in the docking experiments, where the side chains that interact with the ligand the most frequently based on visual analysis are flexible (refer to Materials and Methods: Section 4). Yet again, no discrimination between the receptors was seen (Figure 50).

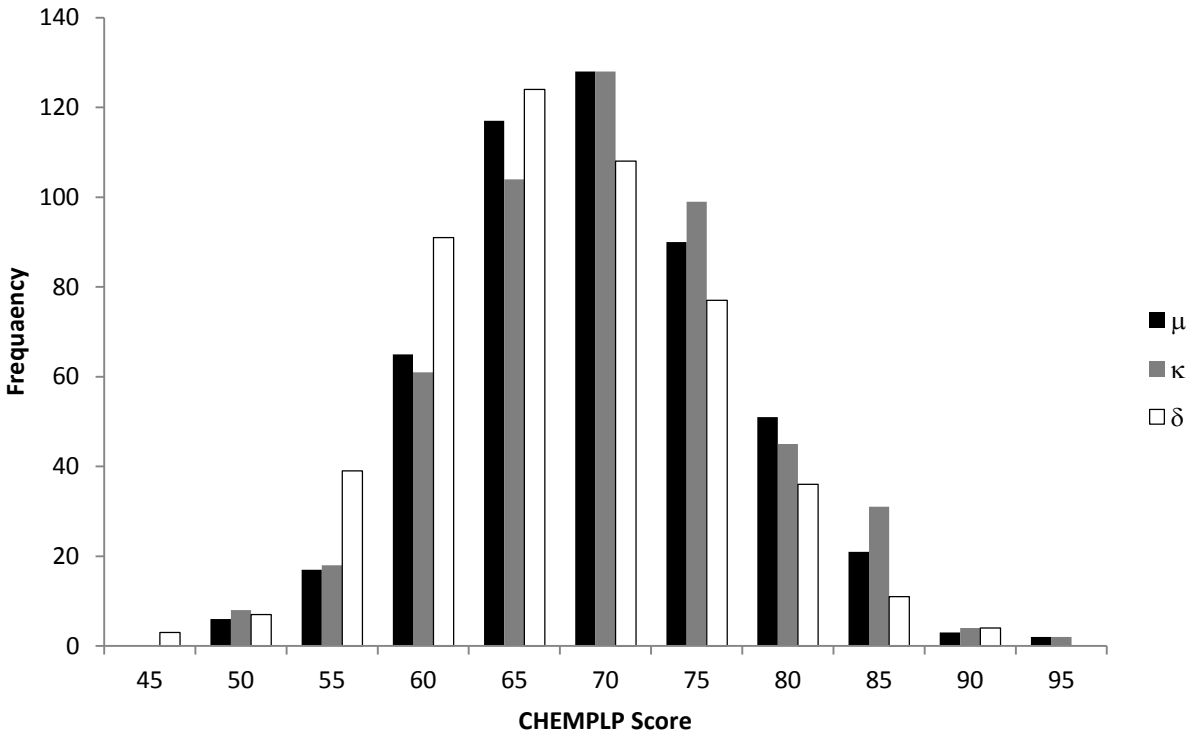



**Figure 50: 500 dockings of DAMGO of the δ, κ and μ receptors with flexible side chains.**

### 5b: Per Amino Acid Scoring Analysis

Since the traditional scoring approach was inadequate for analysing the docking results, the score contribution of individual amino acids was subsequently investigated. GOLD has the ability to display the contribution of each atom of the protein to the docking score. These per atom scores can then be summed up by an awk script run through the Cygwin terminal according to the amino acid to which they belong, providing a score on a per amino acid basis. The advantage of this method is that score contributions from conserved amino acids can be disregarded even if they vary significantly between poses/receptors, as such amino acids would not be responsible for the discrimination seen between the three receptors.

***5bi: μ Receptor Pose Analysis***

Since DAMGO is known to be selective for the μ receptor[49], analysis was focused on the μ receptor docking results. It was found that the top three scoring poses identified as 463, 94 and 75 in the list of 500 results made all the interactions consistent with experimental studies (refer to Introduction: Section 1gi). The reported pose numbers are simply their identification numbers, so their numerical values have no significance.

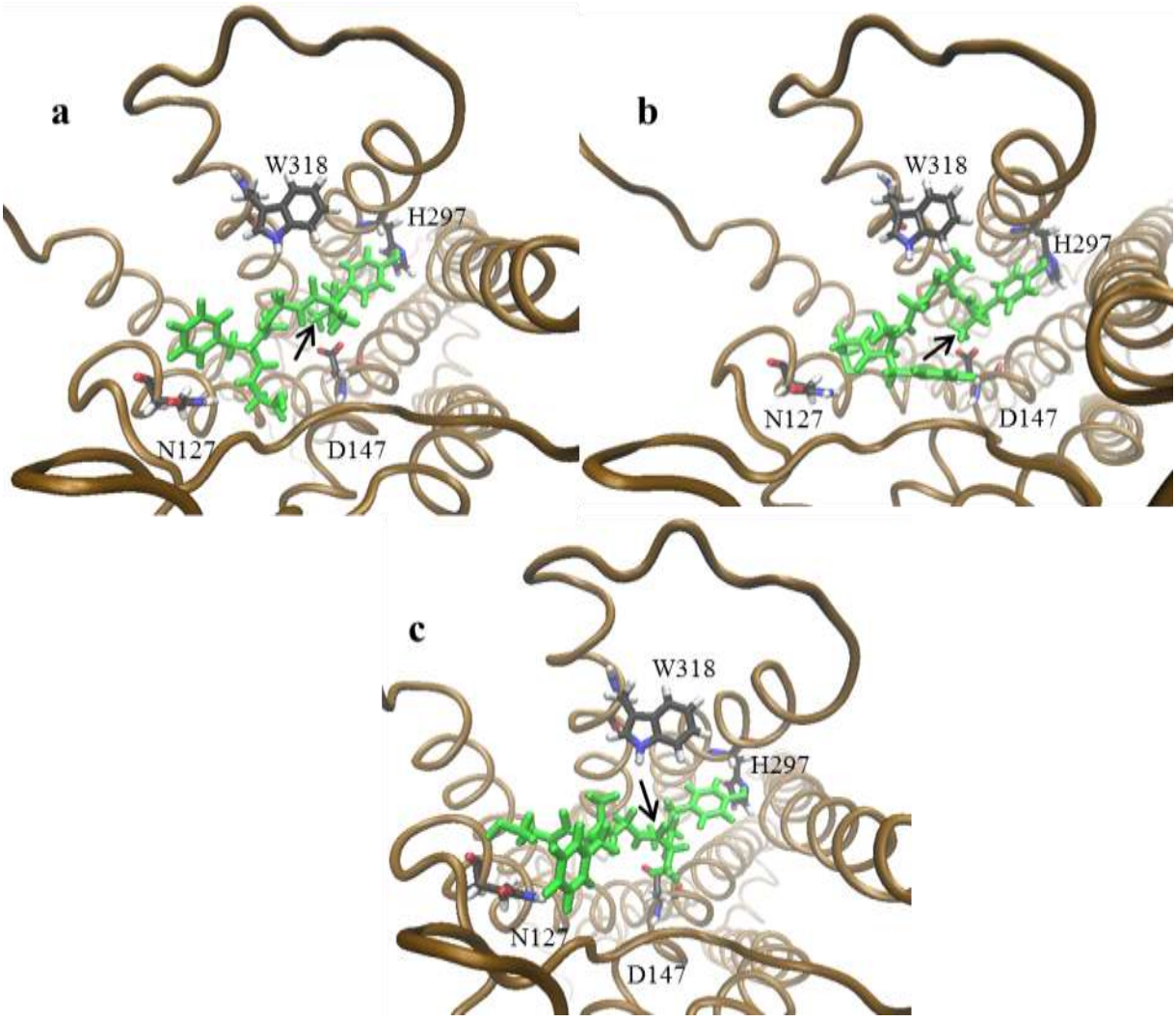



**Figure 51: Visual representation of poses a) 463 (CHEMPLP score = 94.67), b) 94 (CHEMPLP score = 85.56), and c) 75 (CHEMPLP score = 90.87) of DAMGO (green) in the μ receptor. The amino acids N127, D147, H297 and W318 are shown. The black arrow indicates the N-terminus of the DAMGO.**

Several common features were seen among these three poses (Figure 51a-c). A salt bridge between the N-terminus of DAMGO and Asp147 is present in all three poses. Also a H-bond between the phenolic OH of the DAMGO and the imidazole of His297 is found in all poses. The central glycine of DAMGO is found to be in proximity of Trp318. Note that Trp318 is one of the mutations required for DAMGO to bind to the κ receptor (Y312W). Finally, interactions between Asn127 and the N-Me-Phe and Gly-ol groups of DAMGO are found in all the poses. As such, visually (qualitatively) these poses are consistent with what is known experimentally about DAMGO binding to the μ receptor. These poses can also be analyzed quantitatively by scoring the poses on a per amino acid basis.

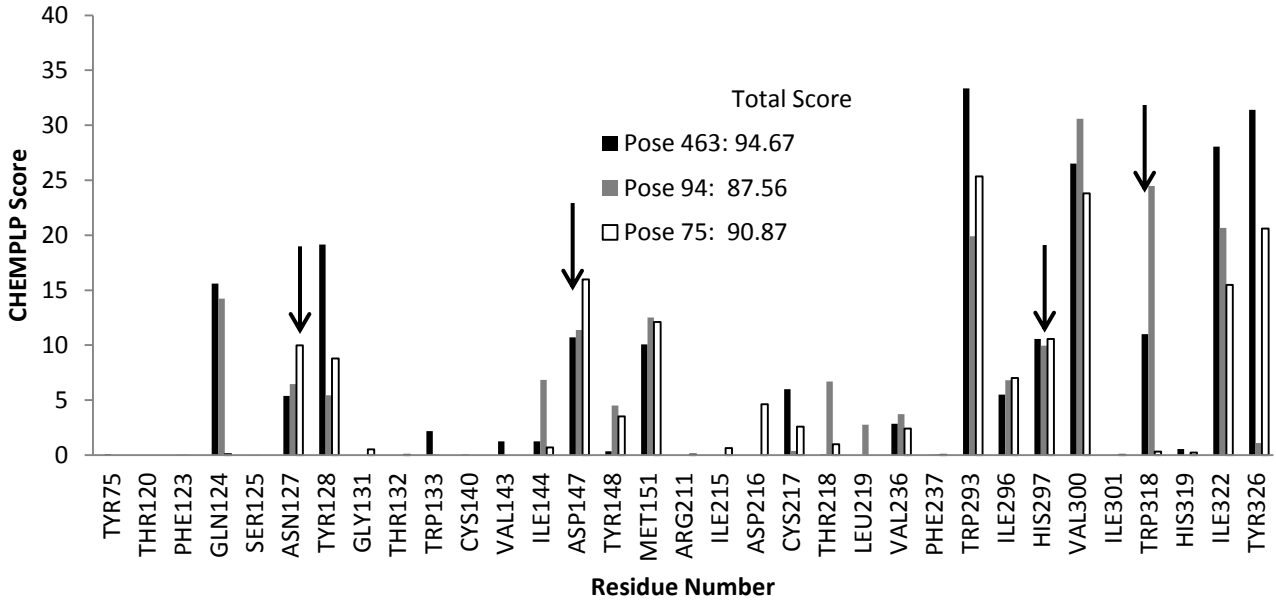



**Figure 52: Individual amino acid contributions to the scores of the DAMGO μ receptor poses 463, 94 and 75.**

Like the total score of a pose, the more positive the score of the amino acid, the more significant the amino acid contribution to the interaction. The scores of all the amino acid residues of a pose, when sums up will equate to the total score of the pose. The four amino acid

residues identified through the qualitative analysis (Asn127, Asp147, His297 and Trp318 indicate by black arrows) contribute to the score of each of the top three docked poses which means that DAMGO is in proximity of these residues. The only exception is that Trp318 makes a minor contribution in pose 75 (Figure 52). There are other amino acid residues that make large contributions to the total score (e.g. Gln124, Tyr128, Met151, Trp293, Val300, Ile322 and Tyr326), however, they are all conserved among the three receptors except Val300.

### *5bii: Pose Rescoring in δ and κ Receptors*

Next, the amino acids determined experimentally to discriminate between the three receptors for DAMGO binding were analyzed among the three top-scoring DAMGO poses. The choice of the top three, rather than four, five or more, poses was chosen arbitrarily, as representing a manageable number for this lengthy analysis.

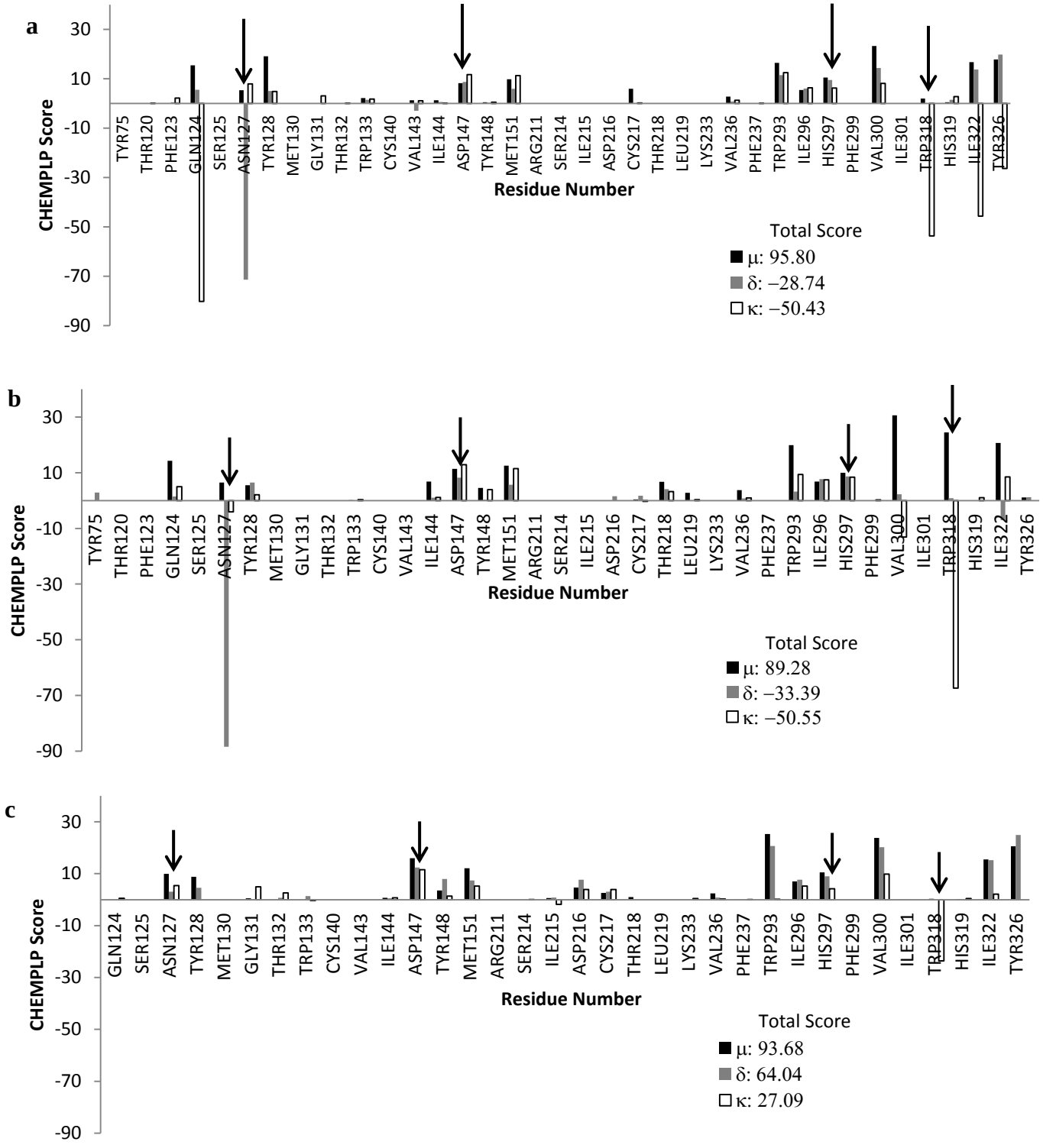



**Figure 53: Per residue scores from the rescoring of DAMGO poses a) 463, b) 94, and c) 75 in the μ, δ and κ receptors.**

The μ receptor poses were rescored back in the μ receptor with rigid (crystal structure) side chains. This operation provides a good reference in analysing the rescoring results of the δ and κ rescores, as only crystal structure side chain conformations can be used in rescoring within different receptors. From the rescoring results of pose 463 (Figure 53a), there was a large negative score in the δ receptor involving Lys108 (at position Asn127), indicating a major steric clash, however there was no major change in the score relative to the μ receptor in the ECLIII (Val300-His319). The κ receptor on the other hand, shows no difference in score at Val118 (at position Asn127) compared with the μ receptor. However, Tyr312 (at position Trp318) had a large negative score, indicating a major steric clash. These two observations are consistent with what is known experimentally about the DAMGO binding. Furthermore, the Y312W is one of the mutations required for DAMGO to bind to the κ receptor. The κ receptor also shows a negative score for the amino acids Gln124, Ile322 and Tyr326, however these are conserved amino acids and the negative scores, which indicate a steric clash, are a result of there being conformational differences in the side chains between the μ and κ receptor. These negative scores were therefore disregarded from further investigation.

A similar pattern was seen in pose 94 (Figure 53b). The δ and κ receptors showed a large negative score at Lys108 and Tyr312, respectively. Compared to pose 463, the major difference was that the contribution from Trp318 (Leu300 in δ) contributed little to the rescoring of pose 463 in the μ and δ receptors, while the score contribution from Trp318 in the μ receptor is high relative to that of the δ receptor in pose 94.

Pose 75 lacks the large negative score at Lys108 (Asn127 in μ) seen in poses 463 and 94. Therefore, this pose was thought not to be a good representation of the actual pose that DAMGO adopts in the μ receptor. Thus it was discarded from further analysis.

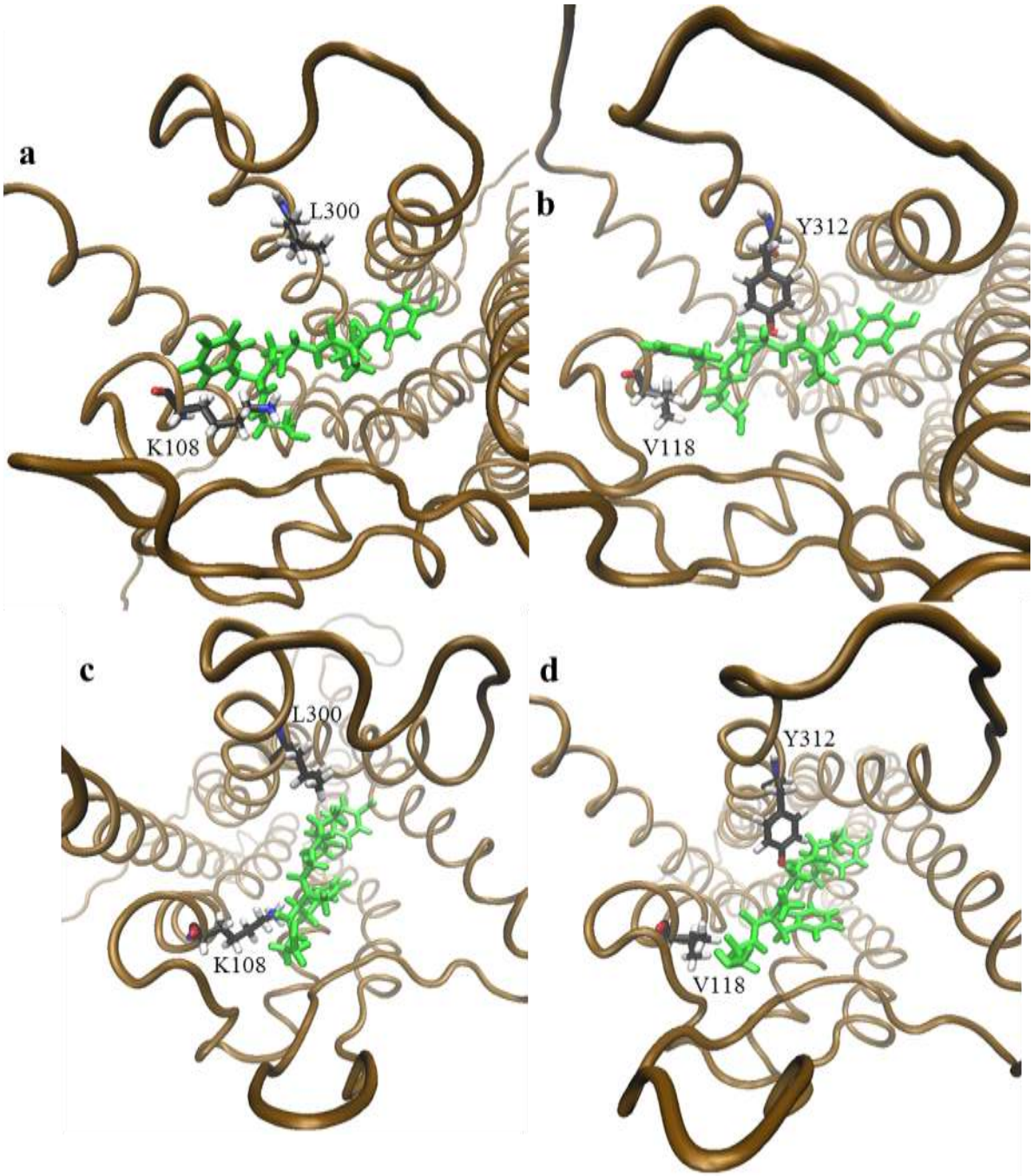



**Figure 54: Visual representations of μ receptor DAMGO (green) poses 463 and 94 rescored in the δ and κ receptors. a) Pose 463 rescored in the δ receptor, b) pose 463 rescored in the κ receptor, c) pose 94 rescored in the receptor, and d) pose 94 rescored in the κ receptor.**

The visual representations of poses 463 and 94 (Figure 54) are consistent with what was seen in Figure 53a,b. A steric clash between Lys108 and DAMGO is present in pose 463 in the δ

receptor, while Leu300 forms only a minor interaction with DAMGO due to the large distance between them (Figure 54a). A hydrophobic interaction between Val118 and DAMGO is found in pose 463 in the κ receptor, while a steric clash is present between DAMGO and Tyr312 (Figure 54b). Pose 94 also exhibits a steric clash between DAMGO and Lys108 when rescored in the δ receptor. No major interaction exists between DAMGO and Leu300 because of the large distance between them (Figure 54c). Pose 94 in the κ receptor engages in a buried interaction with DAMGO via Val118 and the Tyr312 causes a steric clash with DAMGO (Figure 54d).

### 5c: Conclusion Concerning DAMGO Docking

Analysis of DAMGO docking via per amino acid contribution to the CHEMPLP score provided results that are consistent with experimental observations[31,32,55]. This analysis of DAMGO is also the first prediction of the conformation of DAMGO in the μ receptor. Next, this methodology will be used for analysing the β-casomorphin docking results. As there are no mutational or chimeric data on β-casomorphin binding, there are no specific interactions for which to form a baseline comparison.

## 6: Per Amino Acid Scoring Analysis of β-Casomorphin-5

Since the β-casomorphins have been shown to be μ selective[57], the top three scoring μ receptor poses (46, 2 and 268) from the β-casomorphin-5 flexible side chain docking set of 500 poses were used for the scoring analysis on the per amino acid basis (Figure 55).

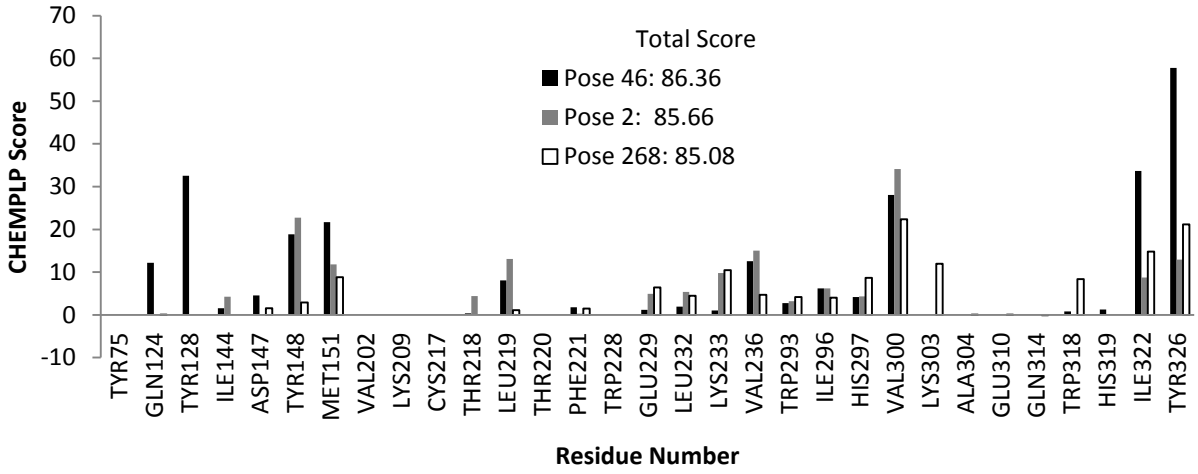



**Figure 55: Individual amino acid contributions of the scores of the β-casomorphin-5 μ receptor poses 46, 2 and 468.**

## 6a: Rescoring of Poses of β-casomorphin-5 from the μ Receptor Docking Experiments in δ and κ Receptors

The three top-scoring μ receptor poses of β-casomorphin-5, 46, 2, and 268, were rescored in the δ and κ receptors. In addition, these poses were rescored back in the μ receptor using rigid side chains, in order to provide an appropriate reference (Figure 56).

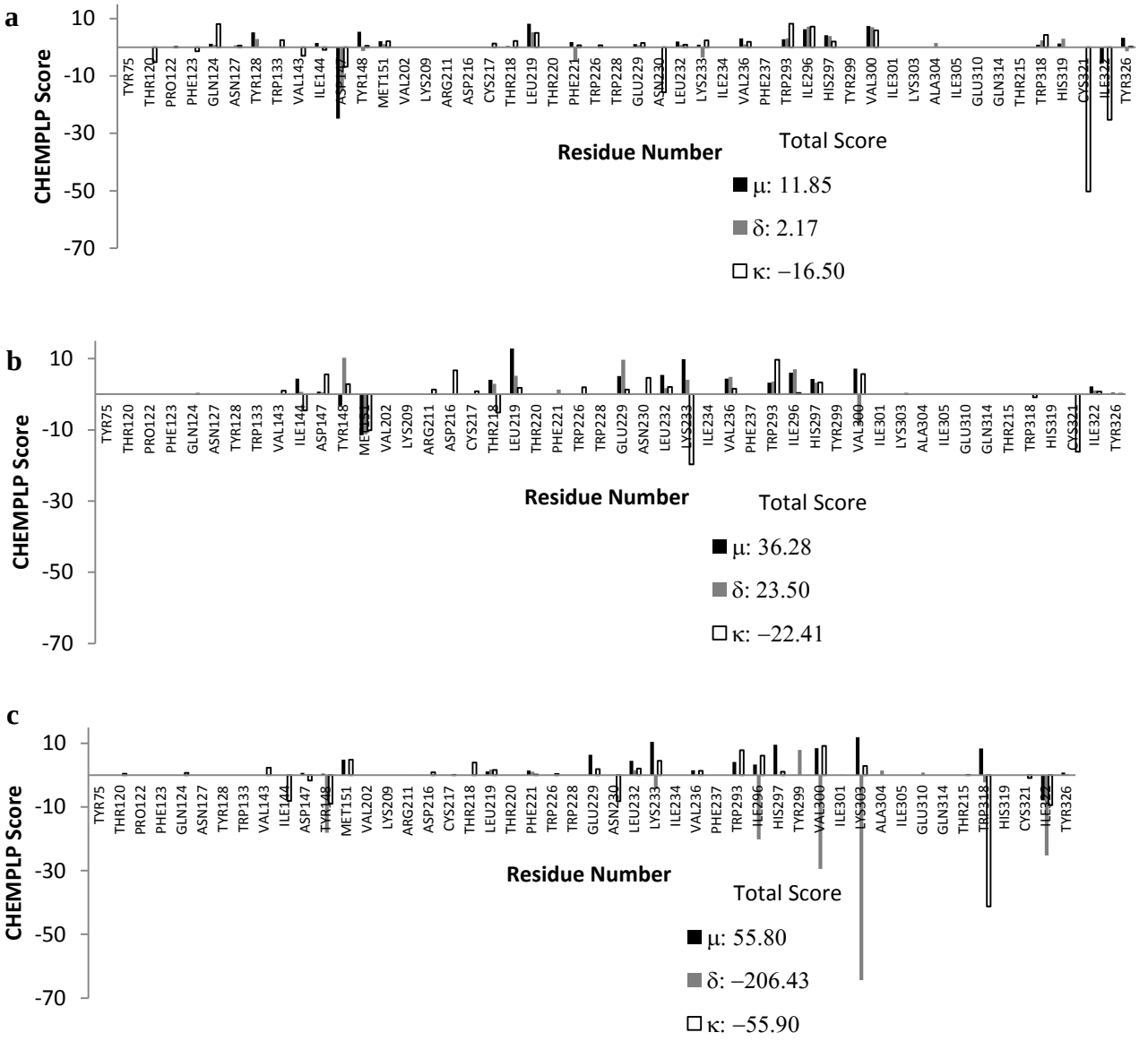



**Figure 56: Residue scores from rescoring of the β-casomorphins μ receptor poses a) 46, b) 2, and c) 268 in the μ, δ and κ receptors.**

The rescoring of pose 46 (Figure 56a) indicated the importance of rescoring the poses back into the μ receptor with its side chains in the crystal structure conformation. For example, Asp147 showed negative scores for the δ and κ receptor rescoring, as well as in the μ receptor. Since the score for this residue in the original scoring with flexible side chains was positive, it

can be assumed that the negative score was a consequence of the side chain conformation. While the same was true for Ile322, Cys321, a conserved residue that was not allowed to be flexible during μ receptor docking but contributed a negative score in the κ receptor rescoring. This was a result of conformational differences in the side chain of the crystal structures. A similar scenario was seen in pose 2 (Figure 56b). While several residues have contributed a negative score relative to the value of the μ receptor rescoring (e.g. Lys233, Val300, Cys321), these residues are all conserved. Both poses 46 and 2 showed no major score difference in any non-conserved amino acid residues. This means that these poses are unlikely to represent to true binding pose of β-casomorphins to the μ receptor. Thus these two poses were discarded from further analysis.

The results from rescoring pose 268 exhibited similar patterns to poses 46 and 2, in that these were variations in the score contribution for conserved amino acids, e.g. Tyr148, Ile196, Lys233, Val300. The variations in score of two non-conserved amino acids were different for pose 268, that is, Lys303 (Trp284 in δ, Glu297 in κ) and Trp318 (Leu300 in δ, Tyr312 in κ). In the μ receptor, Lys303 formed two sets of interactions with β-casomorphin-5, the first interaction was a hydrophobic interaction between the butyl chain of Lys303 and the side chain of the second Pro of β-casomorphin-5. The second was a salt bridge formed between the side chain ammonium centre of Lys303 and the C-terminal carboxylate of the β-casomorphin-5. Trp318, on the other hand, formed a hydrophobic interaction between its six-membered ring and the second Pro side chain of β-casomorphin-5 (Figure 57a).

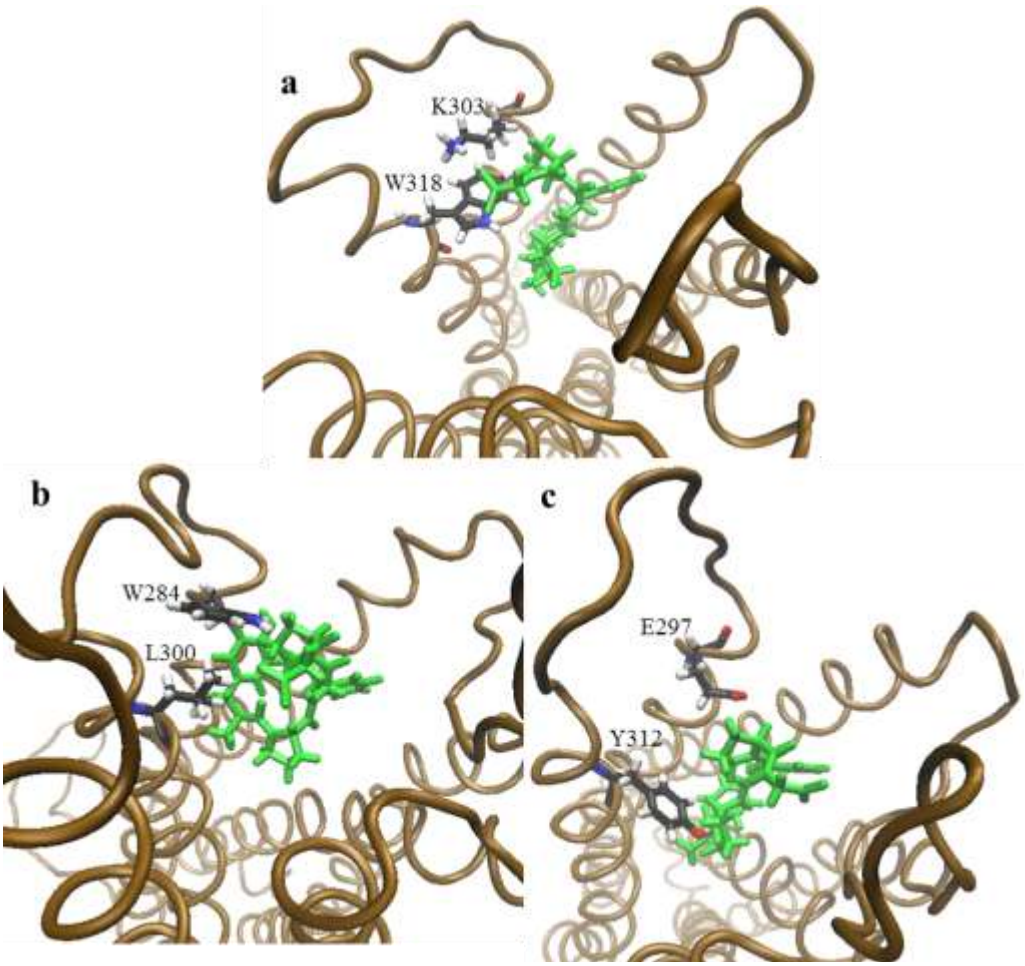



**Figure 57: Visual representations of the rescoring of β-casomorphin-5 pose 268 in the a) μ receptor, b) δ receptor, and c) κ receptor.**

In the δ receptor rescoring, the indole ring of Trp284 (aligned with of Lys303 of the μ receptor) forms a steric clash with the side chain of the second Pro of the β-casomorphin-5. The side chain of Leu300 (aligned with of Trp318 of the μ receptor) forms a minor steric clash with the Tyr $C_\beta$ of the β-casomorphin-5. A buried interaction also occurs between the Leu300 side chain and the N-terminus of the β-casomorphin-5 (Figure 57b). In the κ receptor, Glu297 (aligned with Lys303 of the μ receptor) makes a minor hydrophobic interaction between its ethyl side chain and the second Pro of β-casomorphin-5. The phenolic OH of Tyr312 aligned with

Trp318 of the μ receptor forms a major steric clash with the first Pro side chain of β-casomorphin-5 (Figure 57c).

**6b: Mutant μ Receptor Rescoring**

In order to verify that the differences in scores contributed by interactions between β-casomorphin-5 and the amino acids (Lys303 and Trp318) in the μ receptor and their counterparts in the δ and κ receptors are responsible for the binding selectivity, mutations were introduced in the μ receptor. Lys303 was mutated to Trp (δ) or Glu (κ) while the Trp318 was mutated to Leu (δ) and Tyr (κ). SwissPDB was used to make these point mutations in the receptors. Single mutants were made instead of double mutants in order to examine each of their effects on the total score.

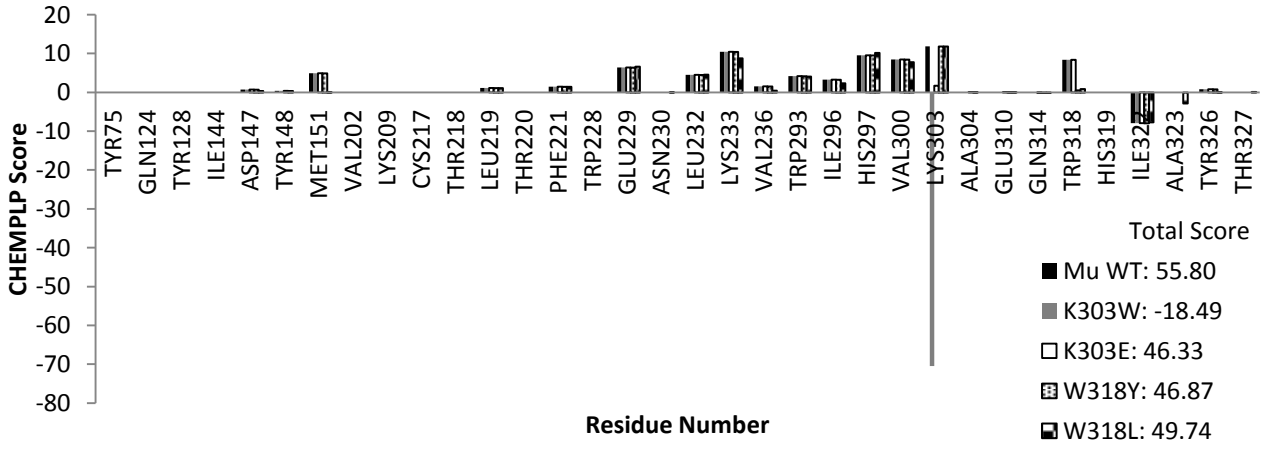



**Figure 58: Per residue scores of the β-casomorphin-5 μ receptor pose 268 rescored in K303 and W318 μ receptor mutants.**

The K303W mutant showed a large negative score at Lys303, which is consistent with its score from the δ receptor rescoring (Figure 56c). The Lys303 for the K303E mutant had a small positive score similar to that seen for the κ receptor rescore of the pose (Figure 56c). The effects of the mutations on Trp318 are slightly different. While the δ and κ receptor rescores were both

negative to varying degrees, they were both marginally positive when rescored in the W318Y and W318L mutants (Figure 58).

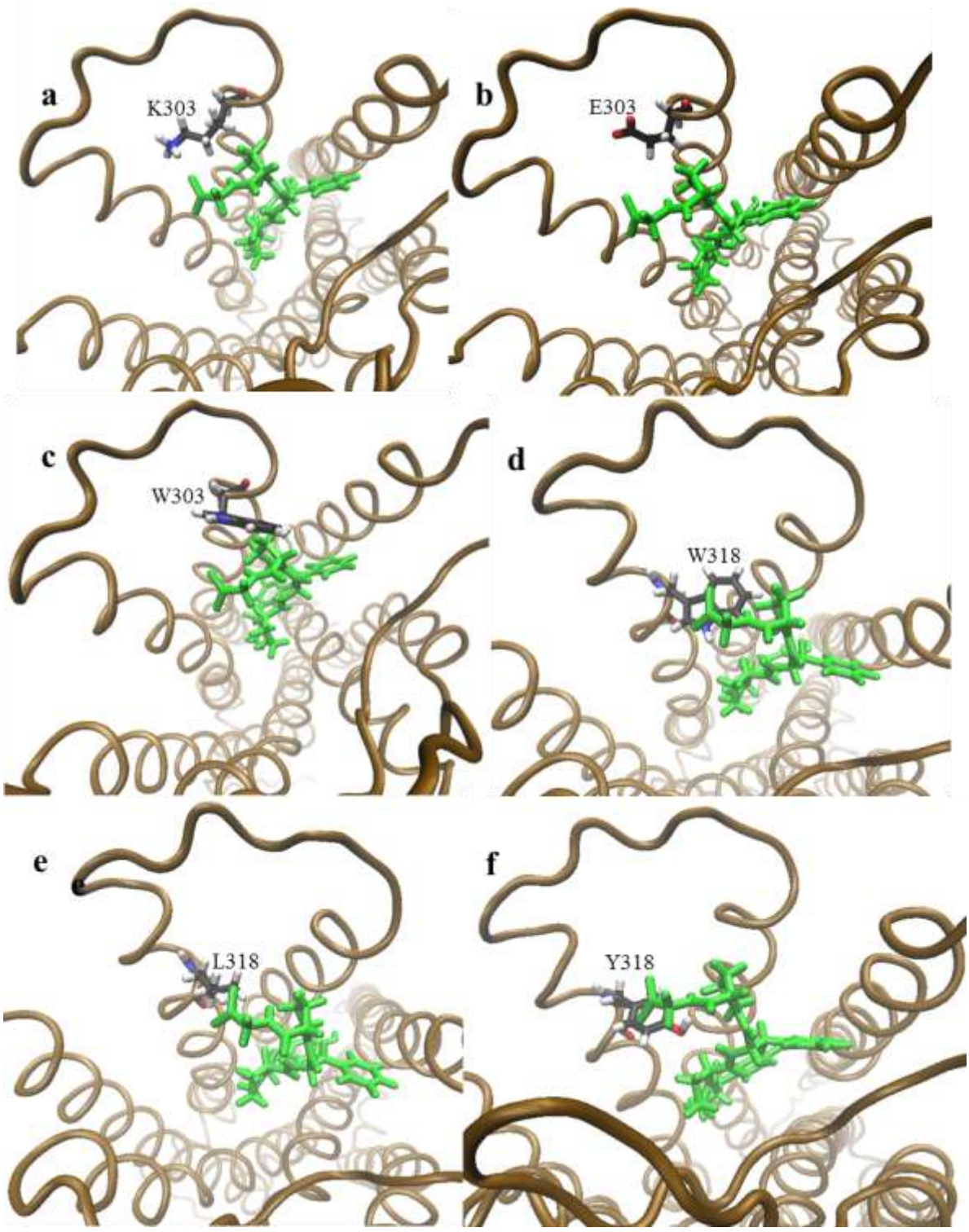



**Figure 59: Visual representations of β-casomorphin-5 μ receptor pose 268 rescored in μ receptor. a) WT (K303), b) K303E, c) K303W, d)WT (W318), e) W318L, and f) W318Y. WT (K303) and WT (W318) are the same μ WT receptor pose with a different amino acid residue displayed.**

As was mentioned previously, Lys303 in the μ WT receptor formed a hydrophobic interaction and a salt bridge with β-casomorphin-5 (Figure 59a). When this residue was mutated to Glu (K303E), only a weak hydrophobic interaction was seen, as seen with the rescoring of the pose in the κ receptor (Figure 59b). When the Lys was mutated to Trp (K303W), the same steric clash occurred as observed in the δ receptor rescoring of the pose (Figure 59c). Trp318 in the WT receptor formed a hydrophobic interaction in β-casomorphin-5 as mentioned previously (Figure 59d). Like the pose rescored in the δ receptor, the W318L mutant makes minimal contact with the β-casomorphin-5 (Figure 59f). The outcome is different for mutation of Trp to Tyr (W318Y). In the rescoring of the pose in the κ receptor, this Tyr residue formed a steric clash via its phenolic OH group. In this W318Y mutant, only a mild steric clash is observed involving the $C_\varepsilon$ of the Tyr but not the OH. This suggests that there is a slight difference in the position of this residue between the κ and μ receptors. It is not surprising that the score is considerable lower than that of the WT μ receptor for this residue (Figure 58).

Based on the mutant rescores, it is hypothesized that the two amino acid residues Lys303 and Trp318 are responsible for the discrimination of β-casomorphin-5 for the μ receptor over the δ and κ receptors.

### 7: Per Amino Acid Scoring Analysis of β-Casomorphin-7 Docked to μ Receptor

Like β-casomorphin-5, β-casomorphin-7 is selective for the μ receptor[57] (Table 4). The results from the docking of β-casomorphin-7 to the μ receptor using flexible side chains (Table 11, Trial 2) were analysed using the per residue scoring analysis employed for β-casomorphin-5. The four top scoring poses of 500 were used (468, 9, 235 and 12). The rationale for choosing the top four, instead of the top three, poses (as used for DAMGO and β-casomorphin-5) is that the additional flexibility derived from the two extra amino acids may make it more difficult to

identify amino acids that discriminate for β-casomorphin-7 binding between the receptors. Thus, as a precaution, an extra pose was analyzed. The scores of these four poses were analysed on a per amino acid basis (Figure 60).

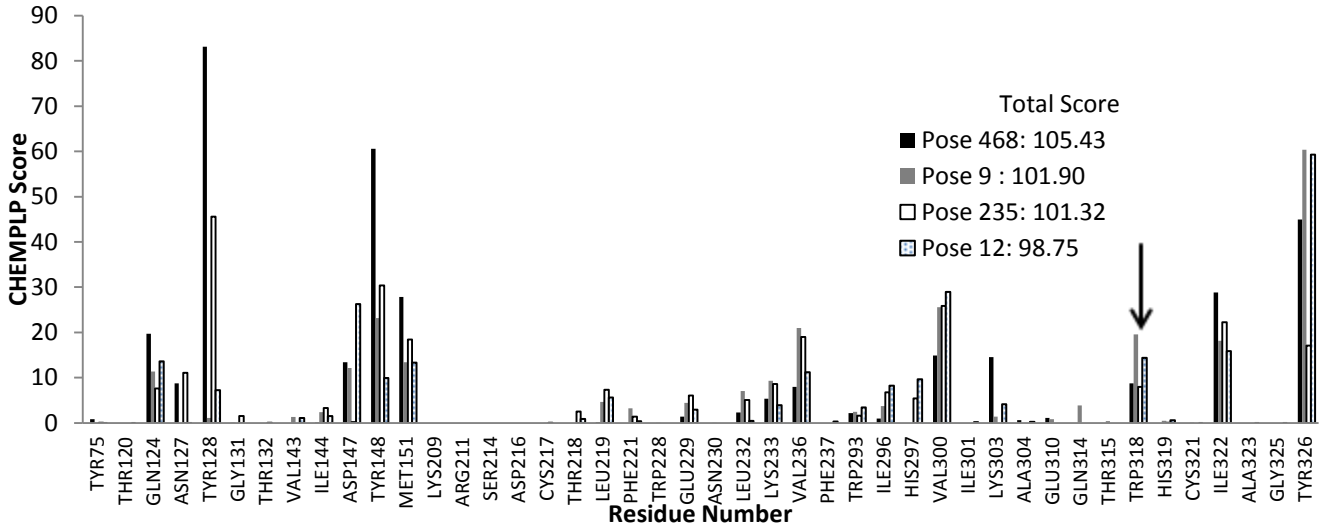
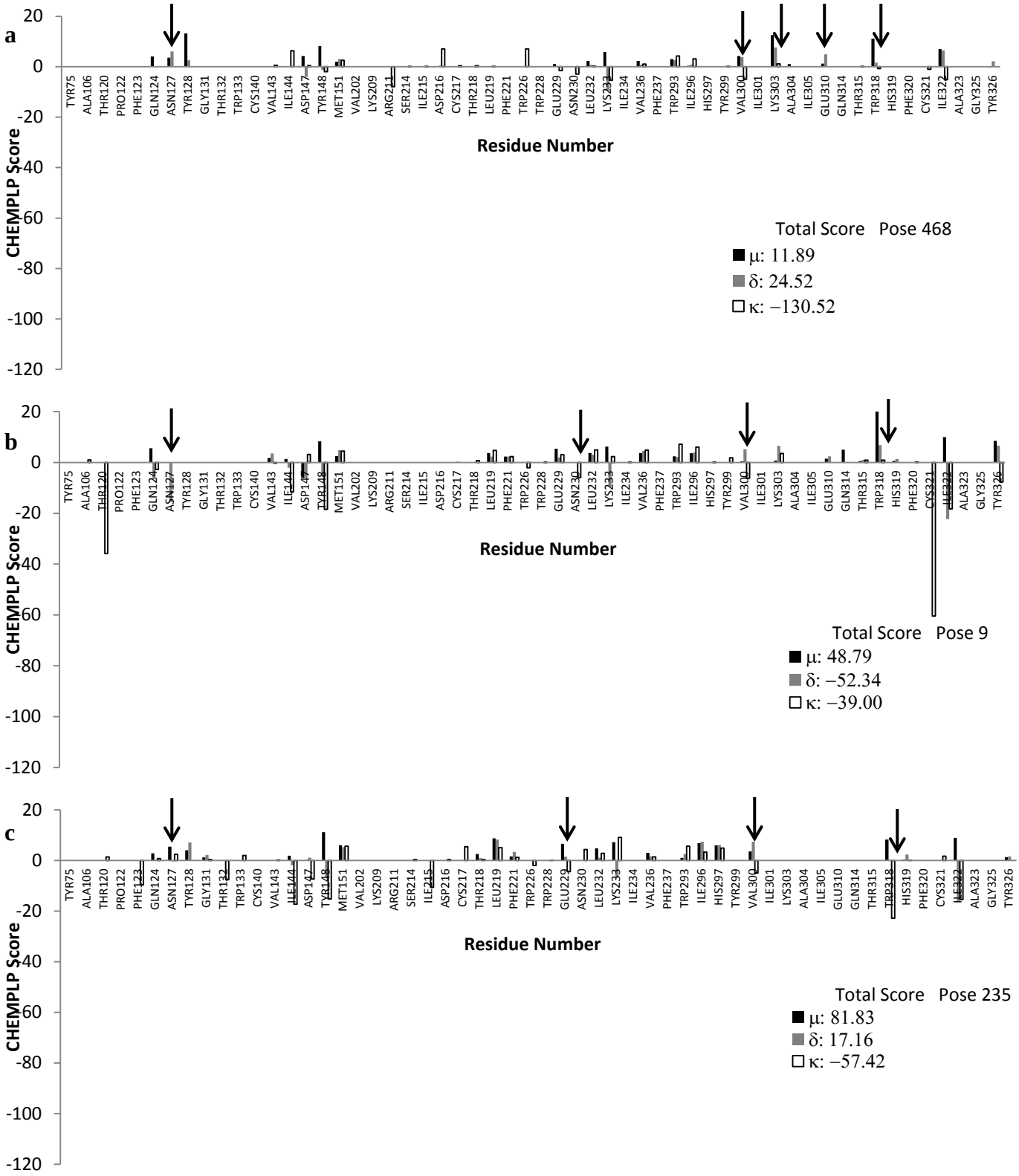



**Figure 60: Individual amino acid contributions towards the scores of the μ receptor poses 468, 9, 235 and 12.**

Many of the conserved amino acid residues that had a high score for β-casomorphin-5 (Gln124, Tyr128, Tyr148, Met151, Val236, Val300, Ile322 and Tyr326) (Figure 55), also have had high scores for β-casomorphin-7 (Figure 60). What is different with β-casomorphin-7 is that non-conserved Trp318 makes a larger contribution to the total score in comparison to β-casomorphin-5. Even though there are many similarities among these three poses in terms of their amino acid residue scores, each of these four poses is geometrically unique.

### 7a: Rescoring μ Poses to δ and κ Receptors

The four poses 468, 9, 235, and 12 were each rescored in the δ and κ receptors. The poses were also rescored back into the μ receptor with rigid side chains from the crystal structure to provide a suitable reference, as was done for the DAMGO and β-casomorphin-5 analyses.


a
b
c
CHEMPLP Score
Residue Number
Total Score Pose 468
μ: 11.89
δ: 24.52
κ: −130.52
Total Score Pose 9
μ: 48.79
δ: −52.34
κ: −39.00
Total Score Pose 235
μ: 81.83
δ: 17.16
κ: −57.42

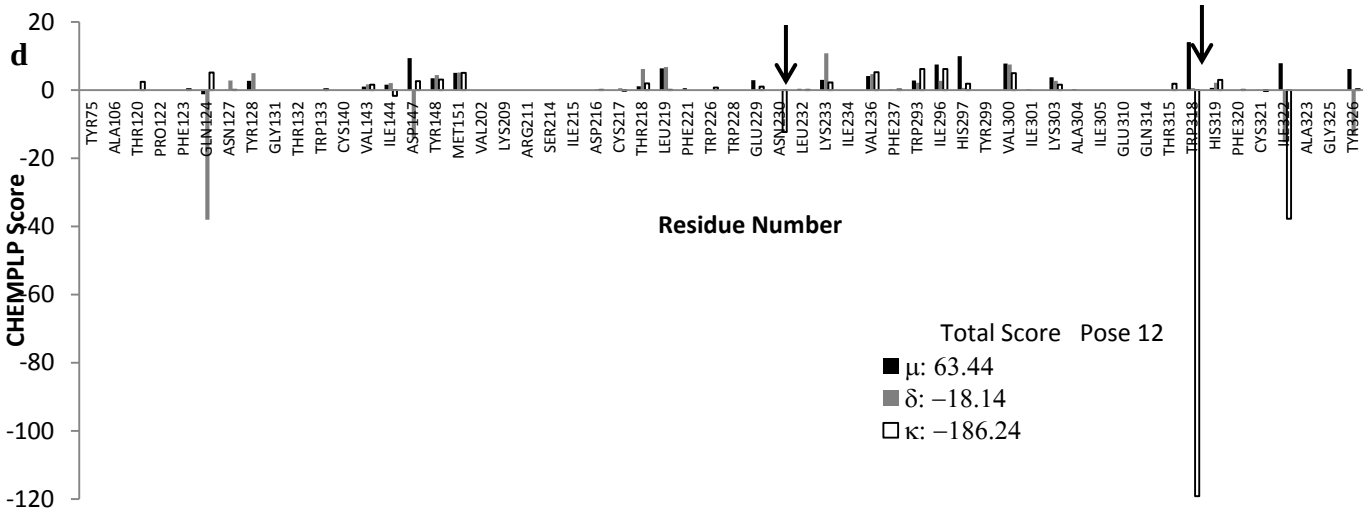



**Figure 61: Residue scores from rescoring of the μ receptor poses a) 468, b) 9, c) 235, and d) 12 in the δ and κ receptors.**

When the four poses were rescored into the μ receptor with rigid side chains, the amino acid residues that were flexible in the original docking experiment had reduced the score relative to the original score with the flexible side chains. The scores of the amino acid residues of the μ receptor that were not flexible in the original docking did not have their score affected by the rescoring.

Pose 468 (Figure 61a) had five non-conserved residues that scored considerably differently among the receptors, Asn127 (δ: Lys108, κ: Val118), Val300 (δ: Val281, κ: Ile294), Lys303 (δ: Trp284, κ: Glu297), Glu310 (δ: Arg291, κ: His304) and Trp318 (δ: Leu300, κ: Tyr312). For Asn127, the score did not change when pose 468 of β-casomorphin-7 was rescored in the δ receptor but the score was zero when rescored in the κ receptor. The Asn127 in the μ receptor forms a buried interaction with its side chain amide and the aromatic ring of the Tyr of the β-casomorphin-7 (Figure 62a). The butyl chain of Lys108 in the δ receptor formed a hydrophobic interaction with the same aromatic ring (Figure 62c). The Val118 of the κ receptor made no interactions with this phenolic group or any other group on the β-casomorphin-7, thus a

zero score was assigned (Figure 62e). Val300 is conserved between the μ and δ receptors and thus no difference is observed in the score of this residue. Val300 makes a hydrophobic contact with the second Pro residue of β-casomorphin-7 (Figure 62a). However, the κ receptor has an Ile at this position, and displays a negative score with β-casomorphin-7 (Figure 61a). The extra methylene group present in the Ile, as compared with Val, faced into the centre of the binding pocket (Figure 62e), resulting in a steric clash with the second Pro residue of β-casomorphin-7. The non-conserved Lys303 showed a similar score between the μ and δ receptors but a lower score for the κ receptor (Figure 61a). The butyl side chain of Lys303 forms a hydrophobic interaction with the third Pro and Ile side chains of the β-casomorphin-7. The ammonium centre of Lys303 forms a salt bridge with the C-terminus of β-casomorphin-7 (Figure 62b). In the δ receptor, the aligned residue (Trp284) formed a buried/hydrophobic interaction between its five-membered ring and the third Pro and Ile side chains (Figure 62d). The smaller Glu297 of the κ receptor at this position made a weak buried interaction with the β-casomorphin-7 and Ile side chain, and thus gave a small score (Figure 62f). Glu310 of the μ receptor showed a positive score for the μ and δ receptors but a zero score for the κ receptor (Figure 61a). It formed a buried interaction between its carboxylate and the side chain of Ile of the β-casomorphin-7 (Figure 62a). In the δ receptor, Arg291 is aligned with Glu310 of the μ receptor. Arg291 also makes a buried interaction between its guanyl group and the Ile side chain (Figure 62c). The κ receptor, on the other hand has His304 which at this position, faces outside the binding pocket, and thus no interactions with β-casomorphin-7 are established (Figure 62e).

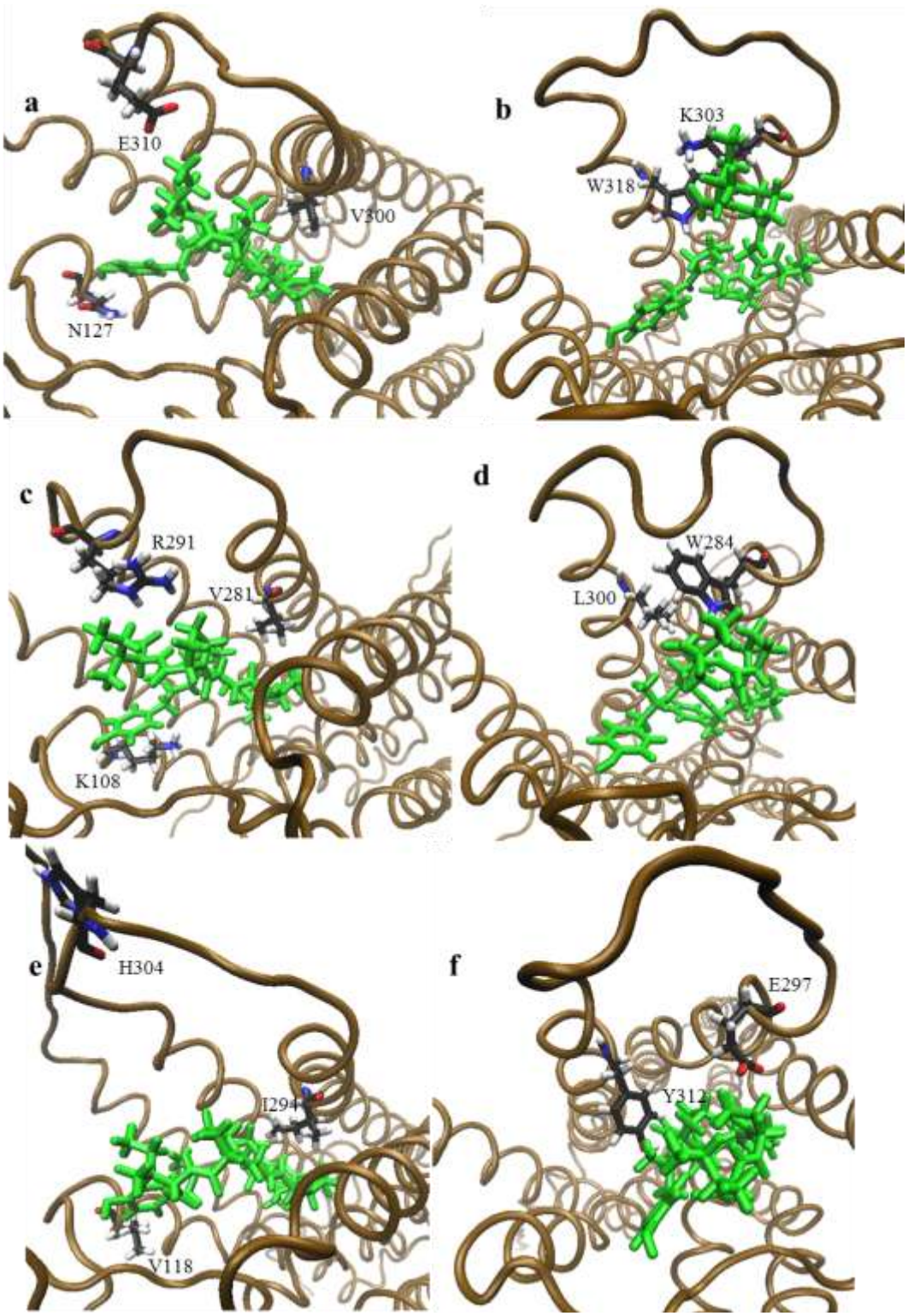



**Figure 62: Visual representation of β-casomorphin-7 pose 468 in the μ receptor rescored in a-b) μ receptor, c-d) δ receptor, and e-f) κ receptor. Each pair of figures from the same receptors is the same rescored pose with different amino acid residues shown.**

The analogous amino acids of the δ and κ receptors to the μ receptor residue Trp318 showed small positive and negative scores for the δ and κ receptors, (Figure 61a). Trp318 of the μ receptor formed a favourable hydrophobic interaction with the third Pro and a buried interaction with the first Pro of the β-casomorphin-7 (Figure 62b). The smaller size of the aligned Leu300 of the δ receptor resulted in a small score due to the relatively large distance between the Leu300 side chain and β-casomorphin-7 (Figure 62c). In the κ receptor, the aligned Tyr312 clashed with the first Pro of the β-casomorphin-7 resulting in a negative score (Figure 62f).

From this analysis of pose 468, Asn127, Val300, Lys303, and Trp318 appear to be to strong candidates for being responsible for receptor discrimination while Glu310 is a weak candidate.

Four non-conserved residues of the three receptors, Asn127 (δ: Lys108, κ: Val118), Asn230 (δ: Thr211, κ: Leu224), Val300 (δ: Val281, κ: Ile294), and Trp318 (δ: Leu300, κ: Tyr312), (Figure 61b) are found to interact differently with Pose 9 of β-casomorphin-7. While Asn127 has a zero score for both the μ and κ receptor rescoring experiments, the aligned Lys108 of the δ receptor has a large negative score. This may suggest a steric clash of the β-casomorphin-7 with the large Lys108 as was seen with DAMGO (Section 5bii). Indeed, in the μ and κ receptors, the respective Asn127 and Val118 residues are too small to make any contact with the β-casomorphin-7 (Figure 63a,e). The Lys108 residue of the δ receptor is so bulky that it sterically clashes with the aromatic ring of the Phe of the β-casomorphin-7 (Figure 63c). Asn230 (μ receptor) has a zero score as does the aligned Thr211 of the δ receptor, but the aligned Leu224 of the κ receptor has a negative score (Figure 61b). This score difference is unexpected as the side chain of μ: Asn230, δ: Thr221 and κ: Leu224 face outside the binding pocket for all three receptors (Figure 63a,d,f). This is a defect in the program as the distance between the clashing

atoms of Leu224 and β-casomorphin-7 were 9.5 Å apart. The Val300 in the μ receptor did not score as high as the Val281 in the δ receptor. There was no visual difference between this Val residue in these two receptors (Figure 63a,c), both making hydrophobic contacts with the first and second Pro side chains of β-casomorphin-7. The larger Ile294 in the κ receptor, however, caused a steric clash with the second Pro of β-casomorphin-7 (Figure 63e). The residues at the equivalent position of Trp318 in the μ receptor showed decreasing scores from μ to δ to κ receptors (Figure 61b). In the μ receptor, Trp318 formed a π-π stacking interaction with the phenolic group of the Tyr and a hydrophobic interaction with the first Pro side chain of β-casomorphin-7 (Figure 63b). In the δ receptor, the Leu300 is engaged in a hydrophobic interaction with the first Pro of β-casomorphin-7 (Figure 63d), while in the κ receptor, a steric clash occurs between the Tyr312 and the second Pro of β-casomorphin-7 (Figure 63f).

From this analysis of pose 9, the amino acid residues Asn127, Val300 and Trp318 are good candidates for amino acid residues that could be responsible for discrimination among the three receptors, while Asn230 is a poor candidate from the analysis.

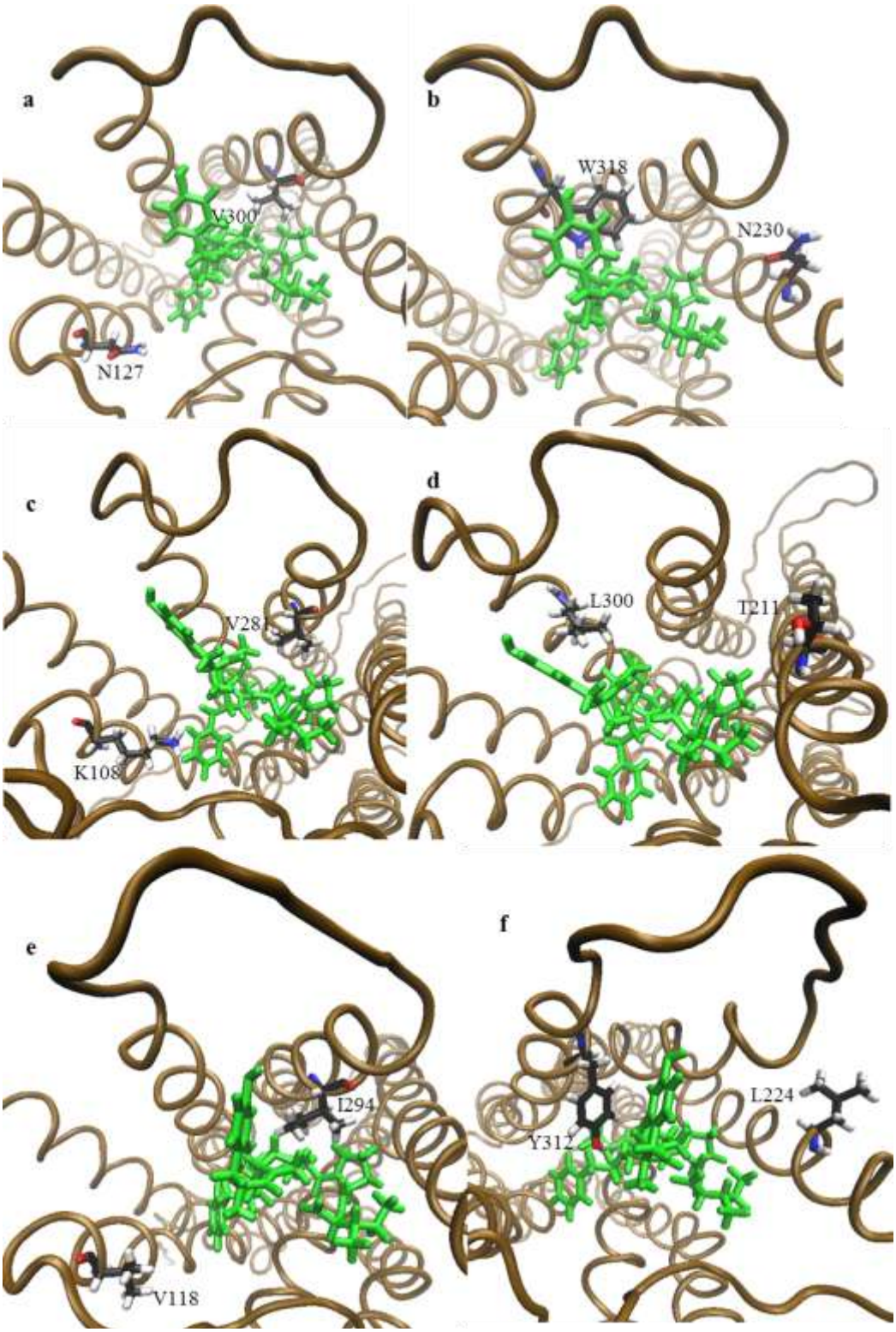



**Figure 63: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a-b) μ receptor, c-d) δ receptor, and e-d) κ receptor. Each pair of figures from the same receptors is the same rescored pose with different amino acid residues shown.**

Four non-conserved amino acids, Asn127 (δ: Lys108, κ: Val118), Glu229 (δ: Asp210, κ: Val223), Val300 (δ: Val281, κ: Ile294) and Trp318 (δ: Leu300, κ: Tyr312), (Figure 61c) have different scores when pose 235 of β-casomorphin-7 is rescored in the three receptors. Asn127 bears a considerably higher score in the μ receptor as compared to those of the δ and κ receptor counterparts. The side chain amide of the Asn127 formed a buried interaction with the side chain of the Ile and an H-bond with the C-terminal carboxylate of β-casomorphin-7 (Figure 64a). The Lys108 of the δ receptor makes a favoured hydrophobic interaction between its butyl chain and the Ile side chain of β-casomorphin-7; however, the ammonium group of Lys108 causes a steric clash with the third Pro of β-casomorphin-7, resulting in an overall low score (Figure 64c). The Val118 of the κ receptor is engaged in a hydrophobic interaction with the Ile side chain of β-casomorphin-7, but not a H-bond, resulting in a low positive score.

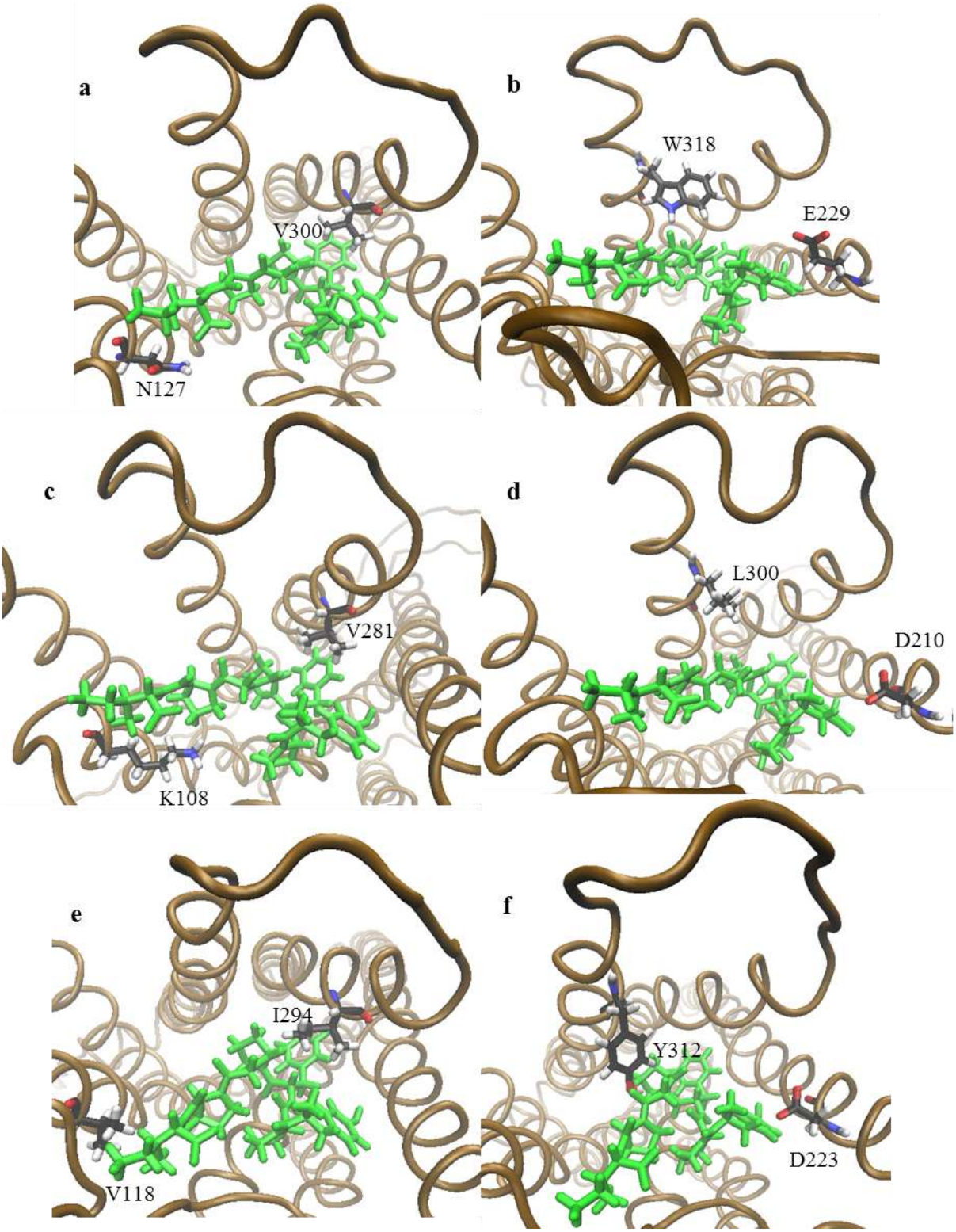



**Figure 64: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a-b) μ receptor, c-d) δ receptor, and e-d) κ receptor. Each pair of figures from the same receptors is the same rescored pose with different amino acid residues shown.**

The counterparts to the Glu229 residue of the μ receptor give lower and negative scores for pose 235 in the δ and κ receptors relative to that of the μ receptor (Figure 61c). In the μ receptor, Glu229 is seen to make a hydrophobic interaction between its ethyl side chain and the $C_\beta$ the Tyr of β-casomorphin-7 (Figure 64b). Both the δ and κ receptors have an Asp residue at this position. In both of these receptors, the carboxylate group of this residue face the $C_\beta$ carbon of the Tyr of β-casomorphin-7, causing a steric clash of varying severity (Figure 64d,f). Pose 235 makes a positive score with Val300 of the μ receptor and the aligned Val281 of the δ receptor but a negative score with the Ile294 in the κ receptor (Figure 61c). The Val300 forms a hydrophobic interaction with the aromatic ring of the Phe of β-casomorphin-7 (Figure 64a), as does Val281 of the δ receptor (Figure 64c). The larger Ile294 of the κ receptor clashes with the second Pro of β-casomorphin-7, thus resulting in a negative score (Figure 64e). Finally, the pose 235 of the β-casomorphin-7 makes a positive score with Trp318 of the μ receptor, a negative score with Tyr312 of the κ receptor and zero score with Leu300 of the δ receptor (Figure 61c). In the μ receptor, Trp318 forms an H-bond between its indole N-H and the carbonyl group of the Gly of β-casomorphin-7 (Figure 64b). The Leu300 in the δ receptor is again too small to form any interactions with β-casomorphin-7, thus it gives zero score (Figure 64d). The Tyr312 of the κ receptor forms a steric clash between its phenolic OH and the second Pro of β-casomorphin-7, resulting in the negative score.

Based on this per amino acid analysis of pose 235, Asn127, Val300 and Trp318 are good candidates for amino acid residues responsible for discrimination between the receptors while Glu229 is a poor candidate for receptor discrimination.

The final pose considered, pose 12 of β-casomorphin-7 (Figure 61d), interacted with only two non-conserved amino acids, Asn230 (δ: Thr221, κ: Leu224) and Trp318 (δ: Leu300, κ:

Tyr312), that discriminate among the receptors. Asn230 in the μ receptor and its δ counterpart Thr211, both had zero score for pose 12. Indeed, neither of these two residues interacts with β-casomorphin-7 (Figure 65a,b). The κ counterpart of this residue, Leu224, for pose 12 (Figure 65c) forms a steric clash with β-casomorphin-7, resulting in a negative residue score (Figure 61d). Trp318 showed a positive score for pose 12 rescored in the μ receptor, a zero score for the δ receptor and a large negative score for the κ receptor. Trp318 formed an H-bond between its indole N-H and the carbonyl group of Phe of β-casomorphin-7 (Figure 65a). Again, Leu300 in the δ receptor makes no contacts with β-casomorphin-7 due to its small size, leading to a score of zero (Figure 65b). The phenolic OH of Tyr312 of the κ receptor stericly clashes with the first Pro residue of β-casomorphin-7 (Figure 65c).

From this per amino acid analysis of pose 12, the amino acid residue Trp318 is a good candidate for a residue that could be responsible for discrimination among the three receptors, while Asn230 is a poor candidate from the analysis.

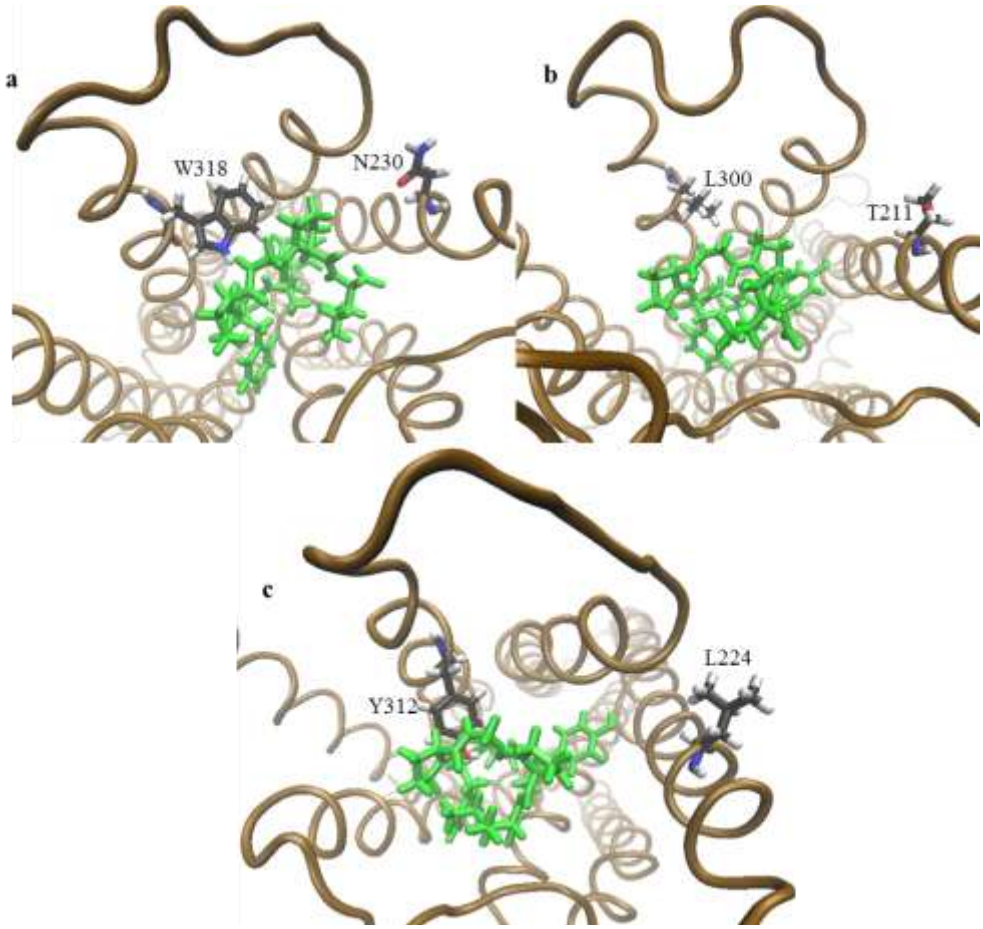



**Figure 65: Visual representation of β-casomorphin-7 (green) pose 12 in the μ receptor rescored in a) μ receptor, b) δ receptor, and c) κ receptor.**

**7b: μ Mutant Rescoring Analysis**

All four top-scoring poses of β-casomorphin-7 selected for rescoring in the μ, δ and κ receptors showed discrimination between several non-conserved residues. This means that none of the poses could be eliminated definitively for consideration as were two of three for β-casomorphin-5 (refer to Results: section 6a). Therefore, μ receptor mutants were made bearing these non-conserved δ and κ amino acid residues at the positions identified previously. The purpose of doing this is to verify that the change in score at the specific amino acid residue is attributed to the difference in the amino acid at that position. Just as with the β-casomorphin-5,

single mutants were made in order to examine each of their effects on the total score, rather than mutating all the amino acid residues for each pose at once.

***Pose 468***

Pose 468 interacted with five non-conserved amino acids, Asn127, Val300, Lys303, Glu310 and Trp318, that showed discrimination between the receptors. As a result, nine mutations were introduced, N127K, N127V, V300I, K303E, K303W, E310H, E310R, W318L and W318Y and pose 468 was rescored in all of them. The WT pose used crystal structure side chain conformations. Note that Val300 is conserved between the μ and δ receptors.

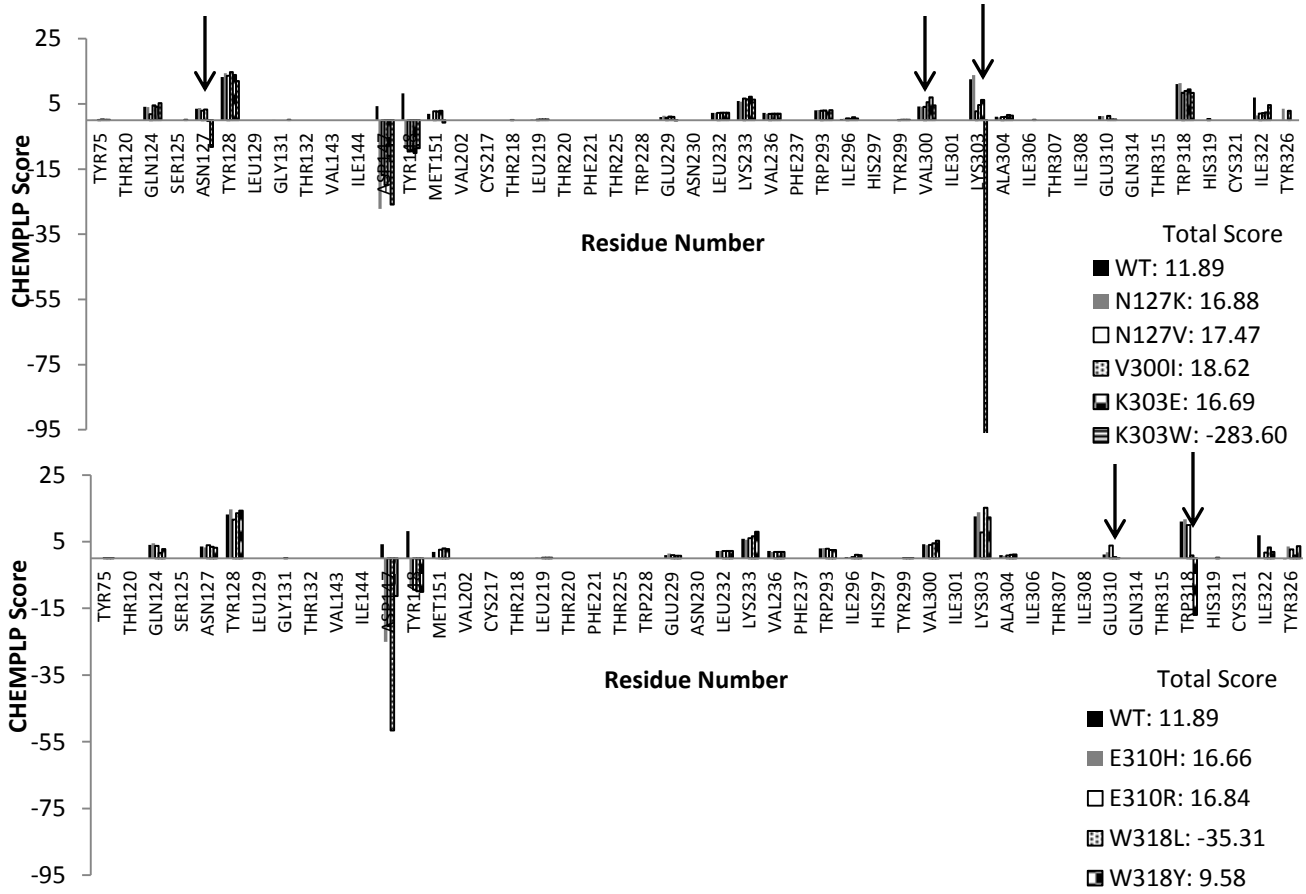



**Figure 66: Residue scores from rescoring of the β-casomorphin-7 μ receptor pose 468 in mutant μ receptors.**

Neither of the N127K and N127V mutants scored differently relative to the WT μ receptor. The same is seen for V300I and K303E mutants. The K303W mutant, however, showed a large negative score relative to that of the WT, indicating a possible steric clash. Neither mutant E310H nor E310R showed any major change in the score, but both Trp318 mutants, W318L and W318Y showed reduced and negative scores respectively (Figure 66).

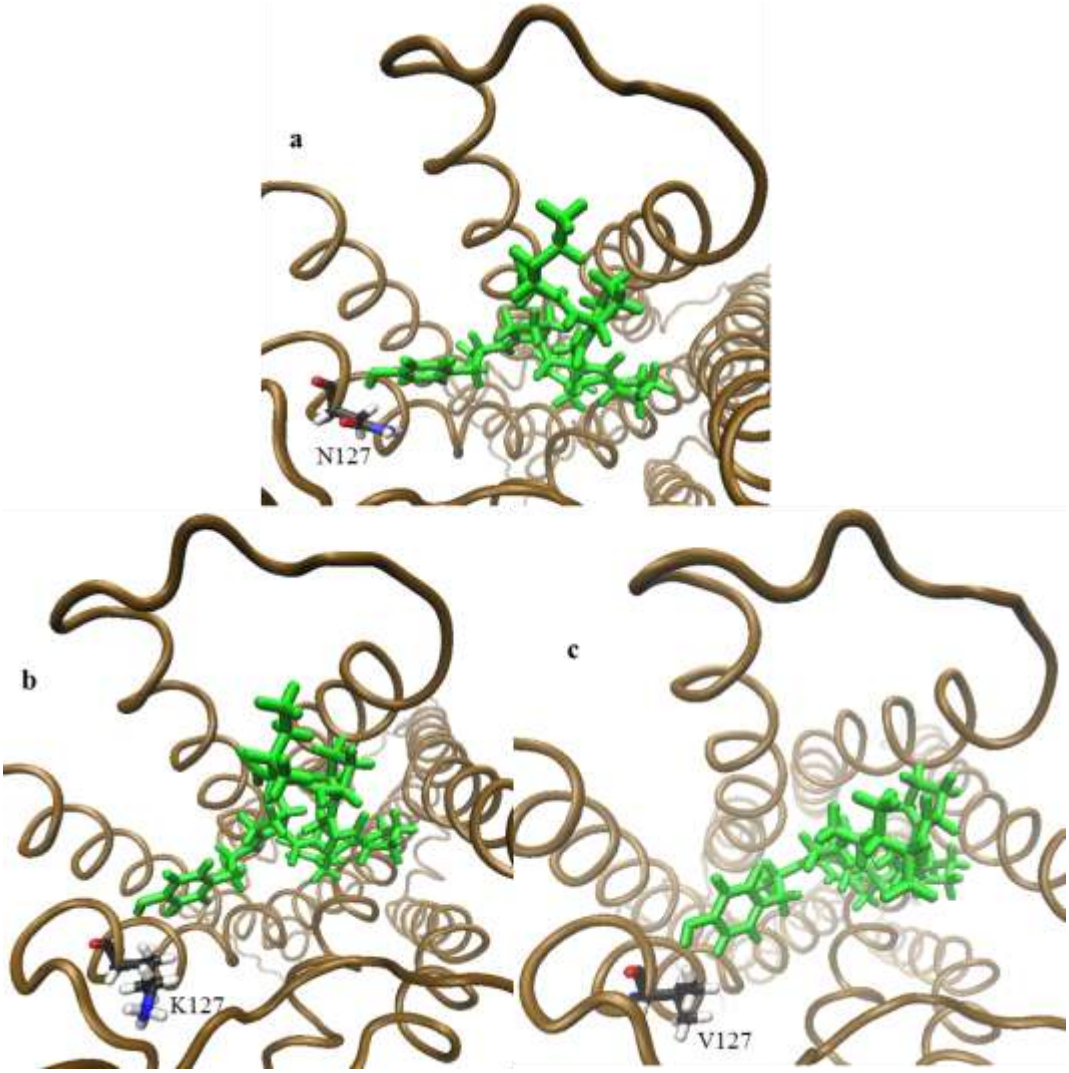



**Figure 67: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) N127K μ receptor, and c) N127V μ receptor.**

Asn127 forms a buried interaction between its side chain amide and the phenolic group of β-casomorphin-7, as mentioned previously (Figure 67a). The N127K mutant also forms a buried interaction between the Lys127 butyl chain and the phenolic OH of β-casomorphin-7 (Figure 67b) while the N127V mutant forms a buried interaction between the Val127 side chain and the phenolic OH of β-casomorphin-7 (Figure 67c). Since all these interactions are similar in strength, no difference in score is observed. It is also important to note that despite the large size of Lys127, it does not cause any steric issues in this pose. Val300 forms a hydrophobic interaction with the second Pro residue of the β-casomorphin-7, mentioned previously (Figure 68a). The V300I mutant, which has an extra methylene group in the Ile side chain, faces away from β-casomorphin-7, thus no difference in score is observed (Figure 68b).

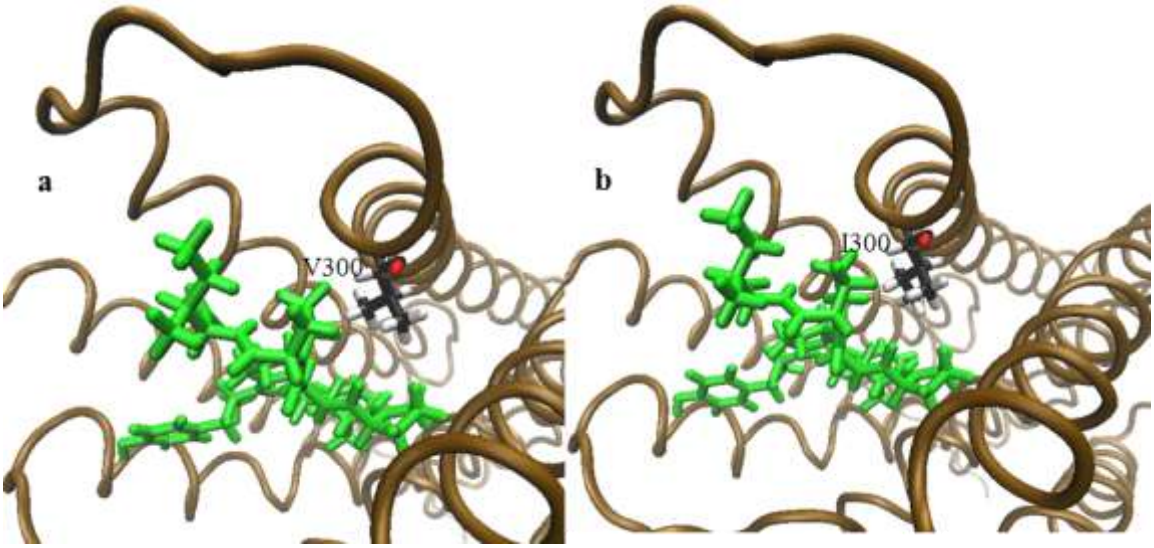



**Figure 68: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor and b) V300I μ receptor.**

As discussed previously, two interactions occur between the Lys303 and the β-casomorphin-7 in the μ receptor. The first is a hydrophobic interaction between Lys butyl chain and the third Pro and Ile side chains of β-casomorphin-7. The second is the salt bridge between

the Lys ammonium center and the C-terminal carboxylate of β-casomorphin-7 (Figure 69a). The K303W mutant formed a large steric clash between the indole ring of the Trp303 and the Ile and third Pro residues of the β-casomorphin-7 (Figure 69b), while the K303E mutant formed a hydrophobic interaction between the ethyl group of the Glu303 and the Ile side chain of β-casomorphin-7 (Figure 69c). Both interactions are consistent with the observations from the δ and κ receptor rescoring for this residue position.

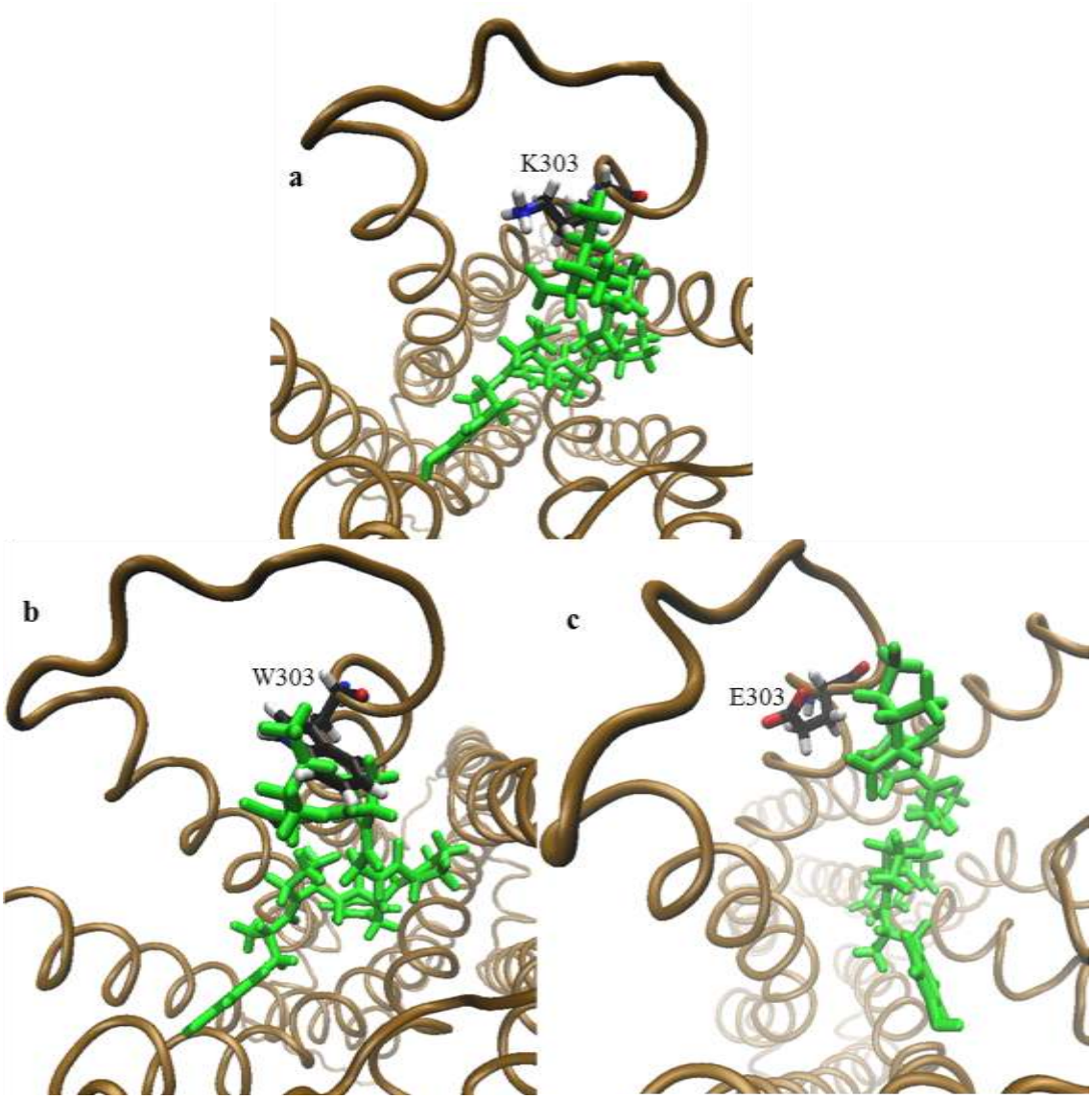



**Figure 69: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) K303W μ receptor, and c) K303E μ receptor.**

The Glu310 carboxylate forms a weak buried interaction with the side chain of the Ile of β-casomorphin-7 (Figure 70a). When the Glu310 is mutated to Arg (E310R), a salt bridge is formed between the guanyl group of the Arg and the C-terminal carboxylate of β-casomorphin-7; however, a steric clash also arises between the guanyl group and the Ile side chain of β-casomorphin-7. Thus these two interactions cancel each other out in terms of their scores (Figure 70b). The E310H mutant forms a weak buried interaction between the imidazole ring of the His310 and the side chain of Ile of β-casomorphin-7. The interaction is weak because of the large distance between the two groups (Figure 70c).

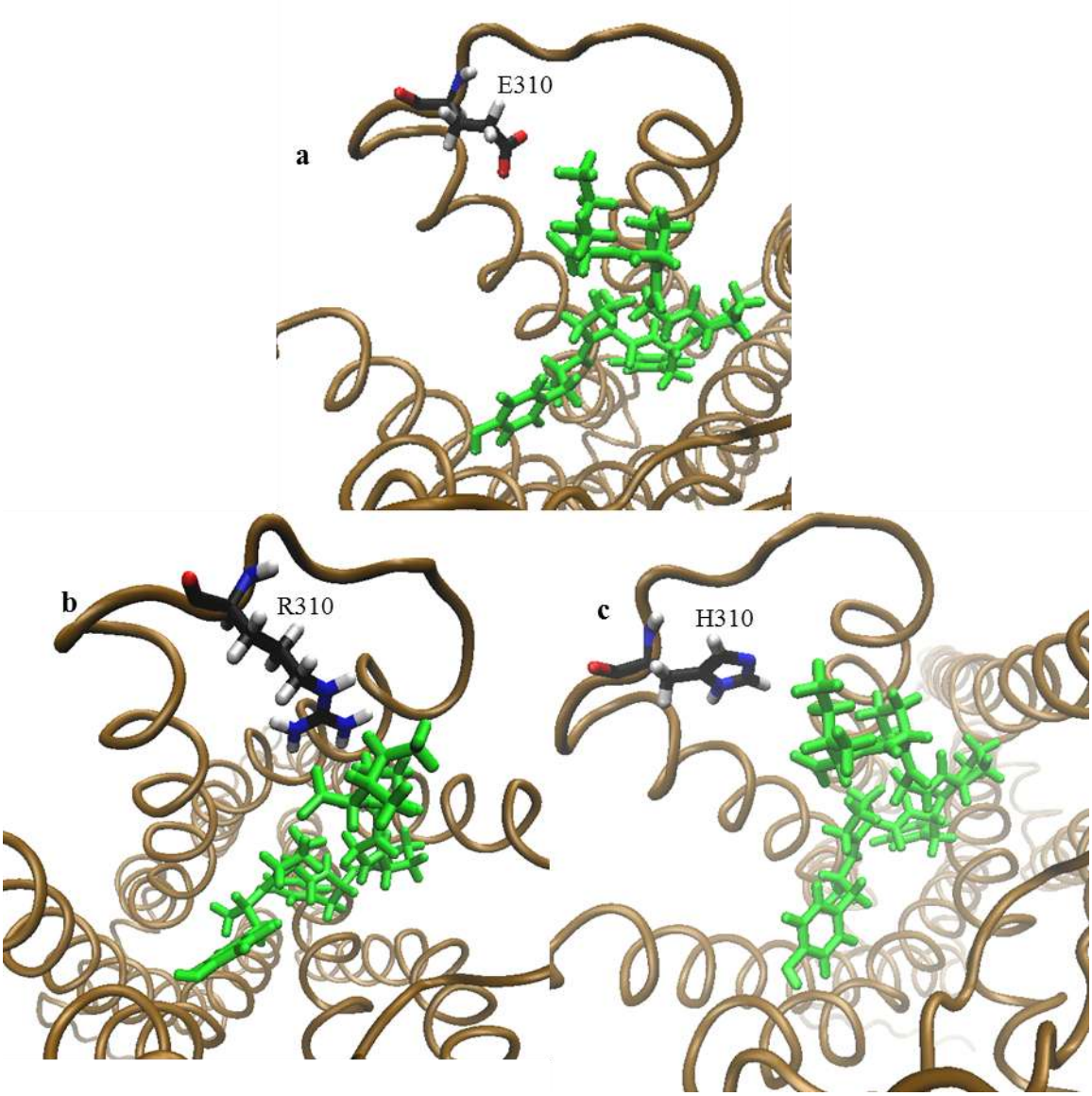



**Figure 70: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) E310R μ receptor, and c) E310H μ receptor.**

Trp318 forms a hydrophobic interaction with the third Pro and a buried interaction with the second Pro side chains of β-casomorphin-7 (Figure 71a). A very weak buried interaction forms between the Leu318 (W318L) and the C-terminal carboxylate of β-casomorphin-7, due to the large distance between the groups (Figure 71b). The W318Y mutant forms a steric clash

between the phenolic OH of the Tyr318 and the side chain of the first Pro of the β-casomorphin-7 (Figure 71c).

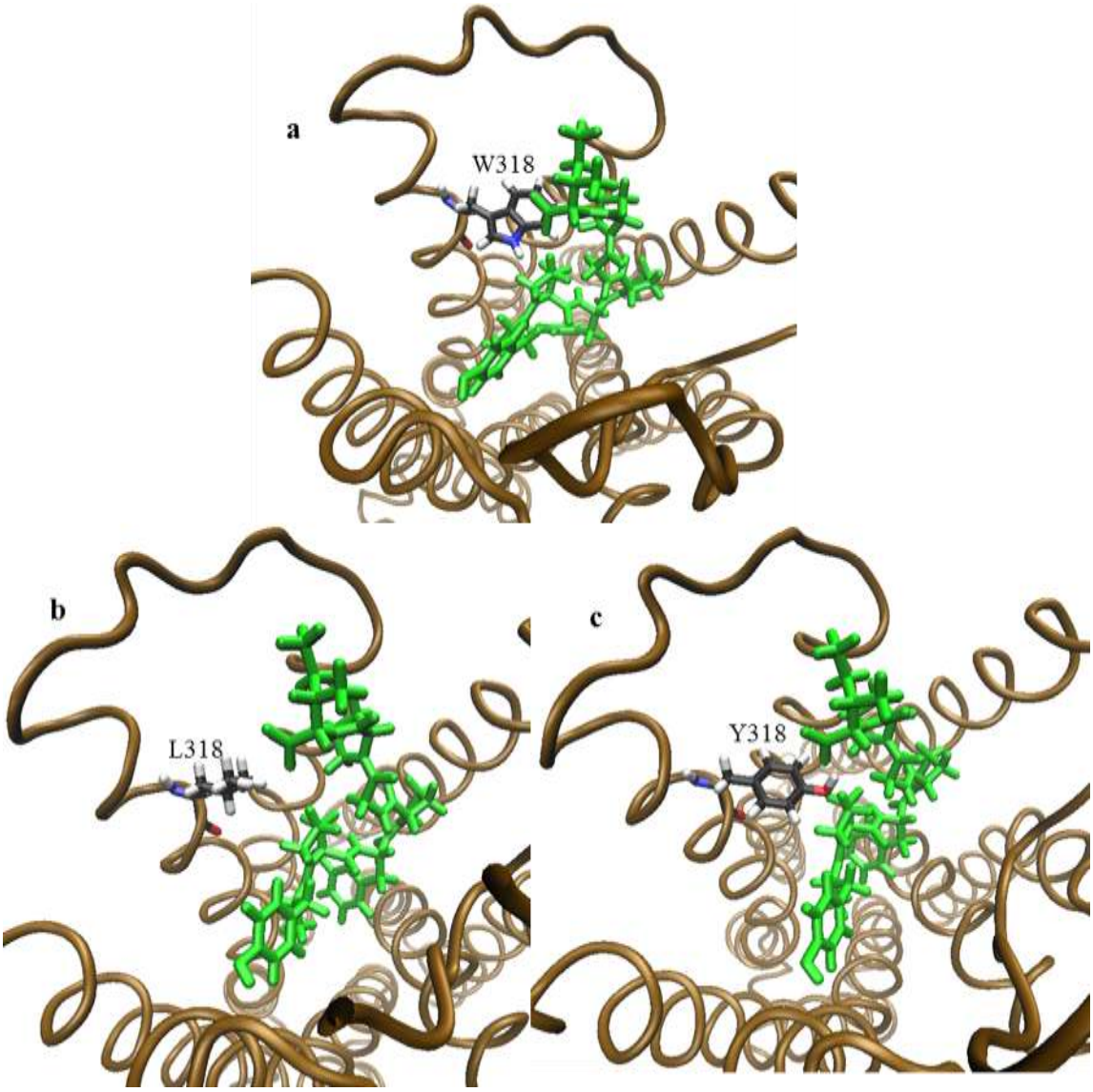



**Figure 71: Visual representation of β-casomorphin-7 (green) pose 468 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor.**

Based on the μ-mutant rescores of pose 468 of β-casomorphin-7, it can be argued that the two amino acid residues Lys303 and Trp318 are responsible for the discrimination of β-casomorphin-7 among the three receptors for pose 468.

***Pose 9***

Pose 9 of β-casomorphin-7 interacted with four non-conserved amino acids, Asn127, Val300, Asn230, and Trp318, that show discrimination between the receptors. Six mutants were introduced into the WT μ receptor, N127K, N127V, N230L, V300I, W318L and W318Y prior to rescoring pose 9. Note that Val300 is conserved between the μ and δ receptors. For residue Asn230, only the κ receptor mutant (N230L) was generated as only the κ receptor showed discrimination at this position. It was found that neither of the Asn127 mutants (N127K and N127V) shows any difference in score relative to the WT μ receptor. The same was true for with the N230L mutant. Mutating the Val300 to Ile (V300I) led to no difference in the score. However, the Trp318 bears a large positive score of the WT Trp. Mutation of Trp318 to Leu and Tyr show a small positive and small negative score, respectively (Figure 72).

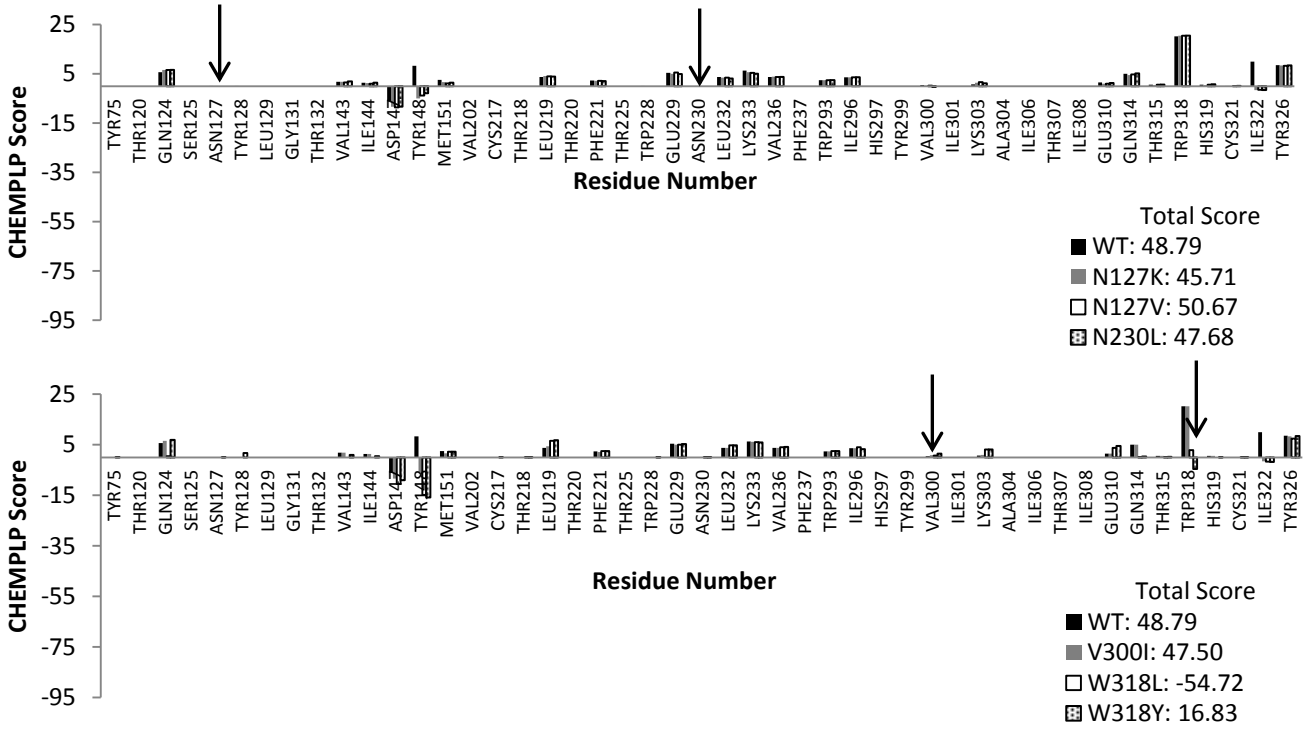



**Figure 72: Residue scores from rescoring of the μ receptor pose 9 in mutant μ receptors.**

As pose 9 for β-casomorphin-7 is positioned rather far from the Asn127 position, none of the amino acids tested (Asn, Lys, Val) makes any contact with it (Figure 73a-c).

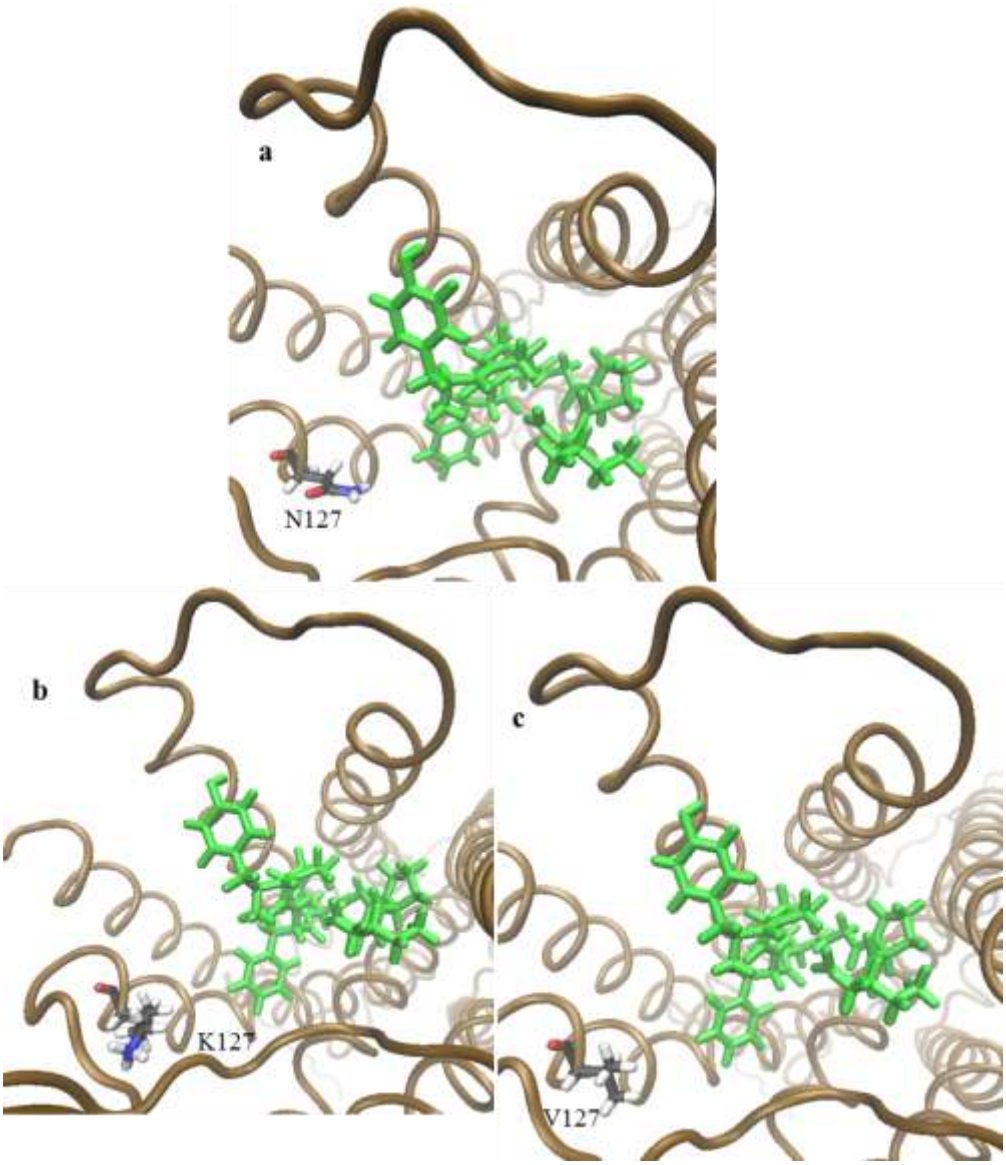



**Figure 73: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor, b) N127K μ receptor, and c) N127V μ receptor.**

Although a negative score was seen in the κ receptor with the Leu224 which is at the position of Asn230 in the μ receptor, no interaction was seen in N230L mutant. This mutant also had a score of zero for this residue (like the WT) (Figure 74b).

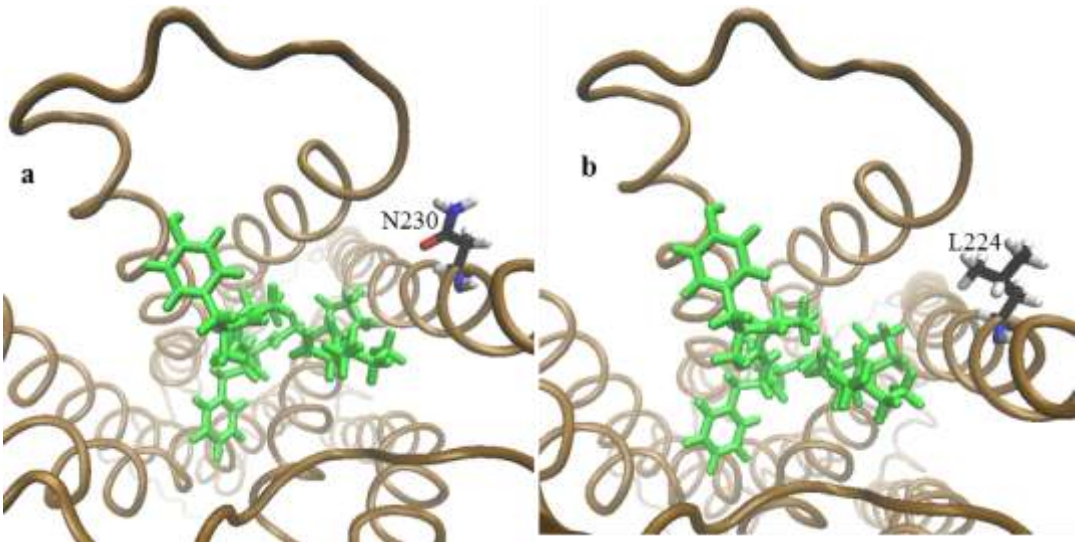



**Figure 74: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor and b) N230L μ receptor.**

The mutation of the Val300 to Ile (V300I) showed no difference in score with pose 9 (Figure 72). The WT Val of the WT μ receptor makes a hydrophobic interaction with the second and third Pro of β-casomorphin-7 (Figure 75a). The I300 makes the same interaction with the extra methylene group facing away from the β-casomorphin-7 (Figure 75b), thus there was no difference in the score.

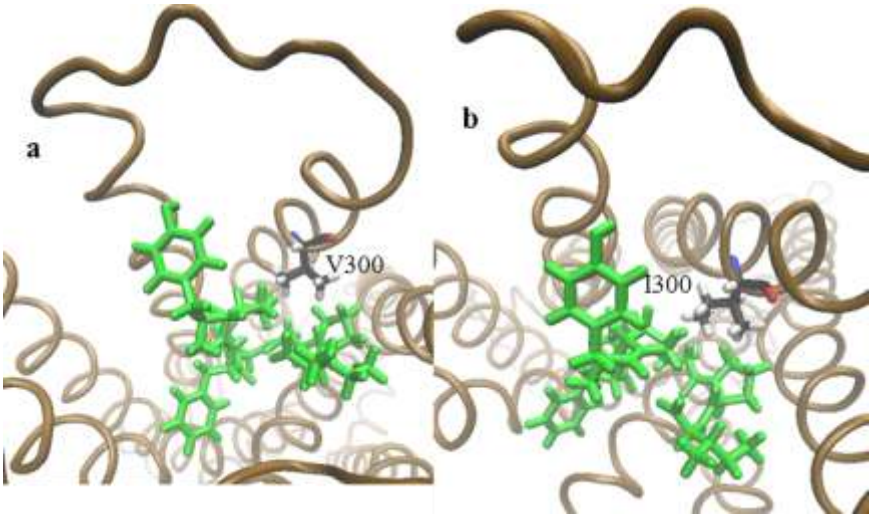



**Figure 75: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor and b) V300I μ receptor.**

Trp318 in the WT μ receptor makes a π-π stacking interaction with the aromatic ring of Tyr and a hydrophobic interaction with the first Pro of β-casomorphin-7, as mentioned previously (Figure 76a). Mutation of Trp to Leu (W318L) resulted in the formation of a hydrophobic interaction between the Leu318 side chain and the aromatic ring of the Tyr of β-casomorphin-7 (Figure 76b). This interaction is not very strong because fewer atoms are involved in the interaction thus it gives a lower score. The score for the W318Y mutant was negative (Figure 72) which is different from that of the κ receptor rescoring (Figure 61b). In the W318Y mutant, Tyr318 is engaged in a much more severe steric clash with the first Pro side chain of the β-casomorphin-7 relative to what is seen in the κ receptor rescoring (Figure 76c).

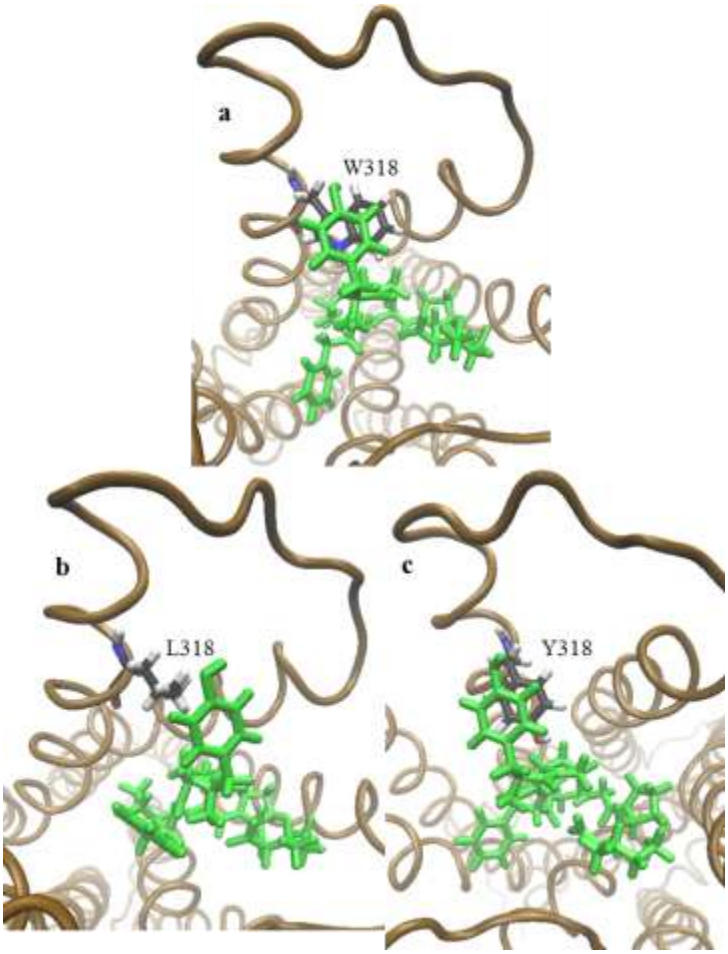



**Figure 76: Visual representation of β-casomorphin-7 (green) pose 9 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor.**

Based on the mutant rescores, it was concluded that the amino acid residue Trp318 is responsible for the discrimination of β-casomorphin-7 between the three receptors for pose 9.

***Pose 235***

Pose 235 of β-casomorphin-7 interacts with four non-conserved amino acids, Asn127, Glu229, Val300 and Trp318, that showed discrimination between the receptors. Six mutations, N127K, N127V, E229D, V300I, W318L and W318Y, were introduced into the WT μ receptor. Again, Val300 is conserved between the μ and δ receptors and both the δ and κ receptors have Asp at the position of E229.

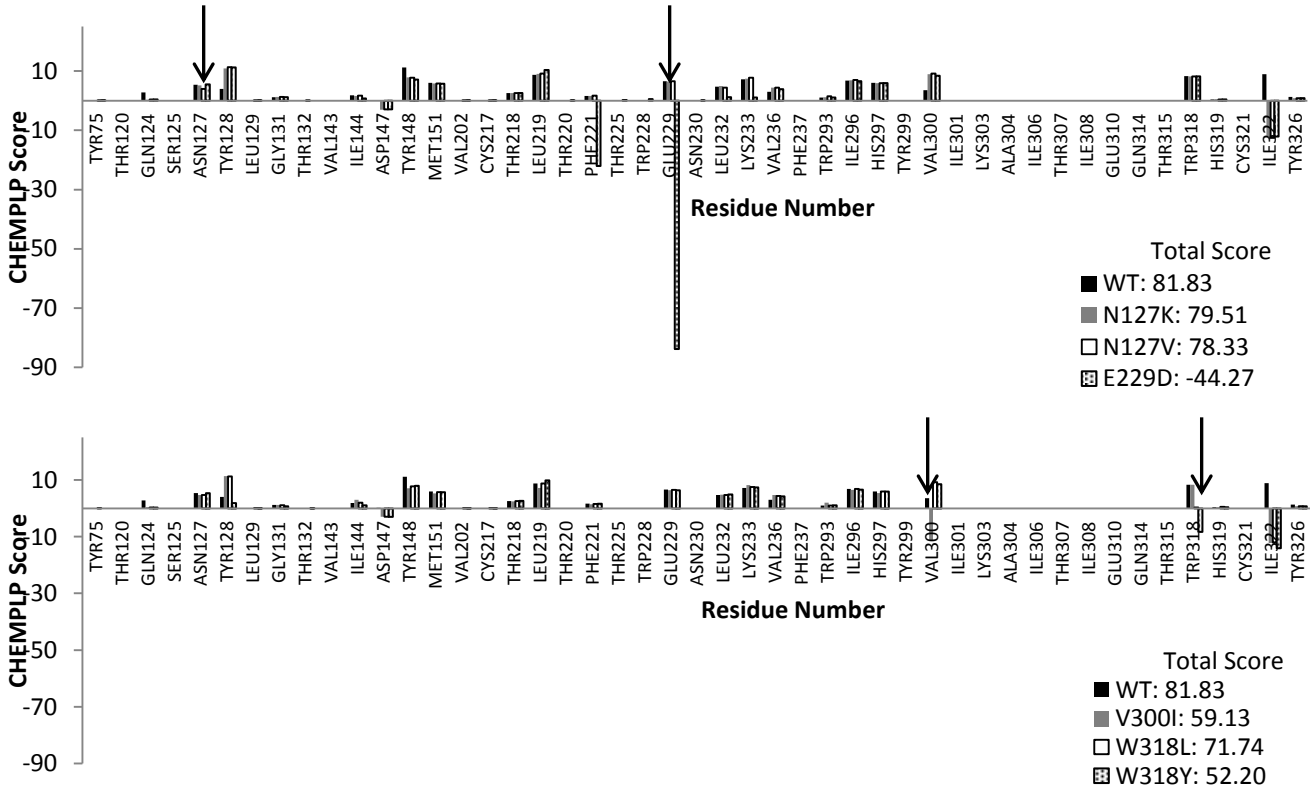



**Figure 77: Residue scores from rescoring of the μ receptor pose 235 in mutant μ receptors.**

It was seen again that no difference in score was found when pose 235 was rescored in the Asn127 mutants of the μ receptor (N127K and N127V). Mutating the Glu229 to the smaller Asp has resulted in a large negative score suggesting occurrence of a steric clash. The V300I mutation that made no difference in score in the previous two poses (468 and 9) led to a negative score for pose 235. The Trp318 experiences the same trend as with the previous poses. The W318L mutant has a very low score (zero in this case) relative to the WT Trp318 and the W318Y mutant has a negative score (Figure 77).

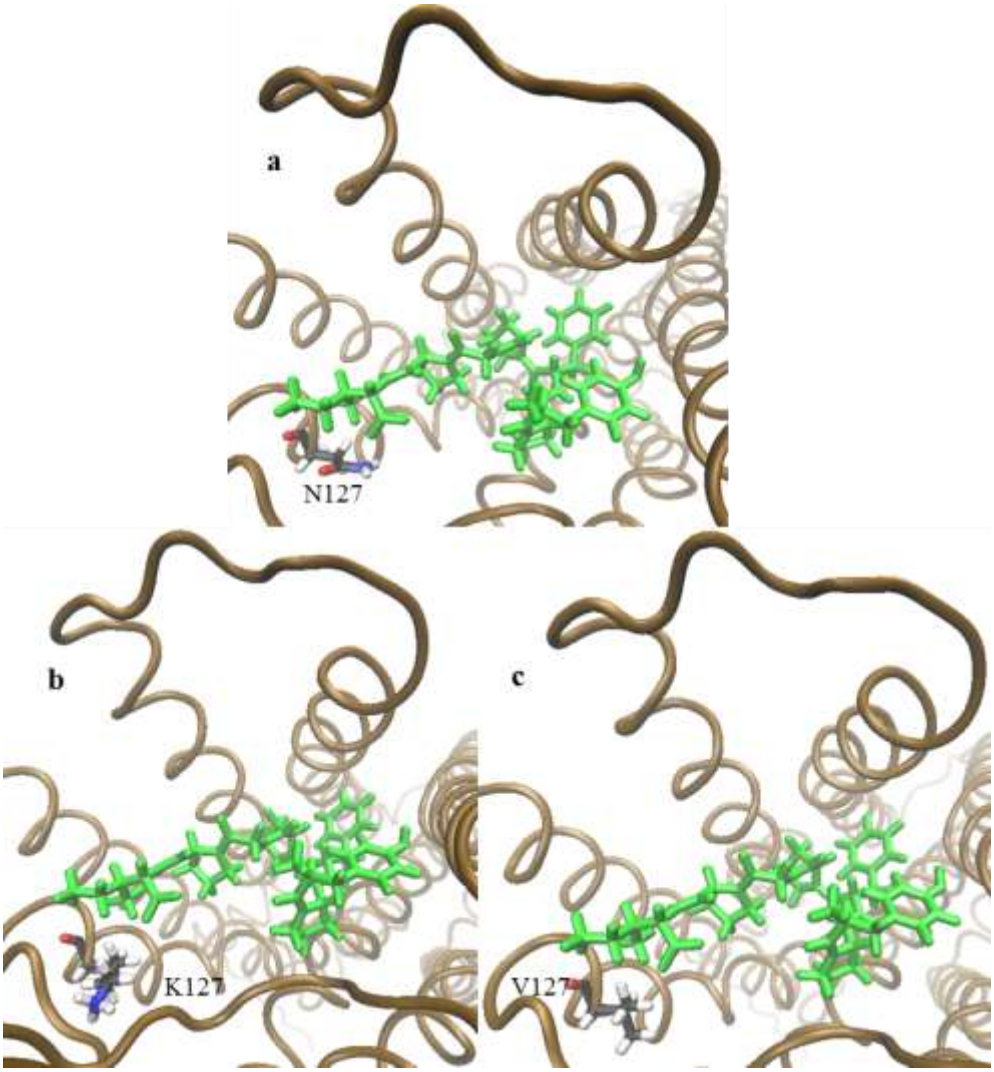



**Figure 78: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor, b) N127K μ receptor, and c) N127V μ receptor.**

The Asn127 of the WT μ receptor forms a buried interaction with the side chain of the Ile and an H-bond with the C-terminal carboxylate of pose 235 of β-casomorphin-7 (Figure 78a). Mutating the Asn to a Lys (N127K) results in the formation of a hydrophobic interaction between the butyl group of the Lys and the side chain of Ile of β-casomorphin-7 (Figure 78b). A similar hydrophobic interaction arises in the N127V mutant (Figure 78c). No score difference was observed between the different residues at position 127 in the mutated μ receptor.

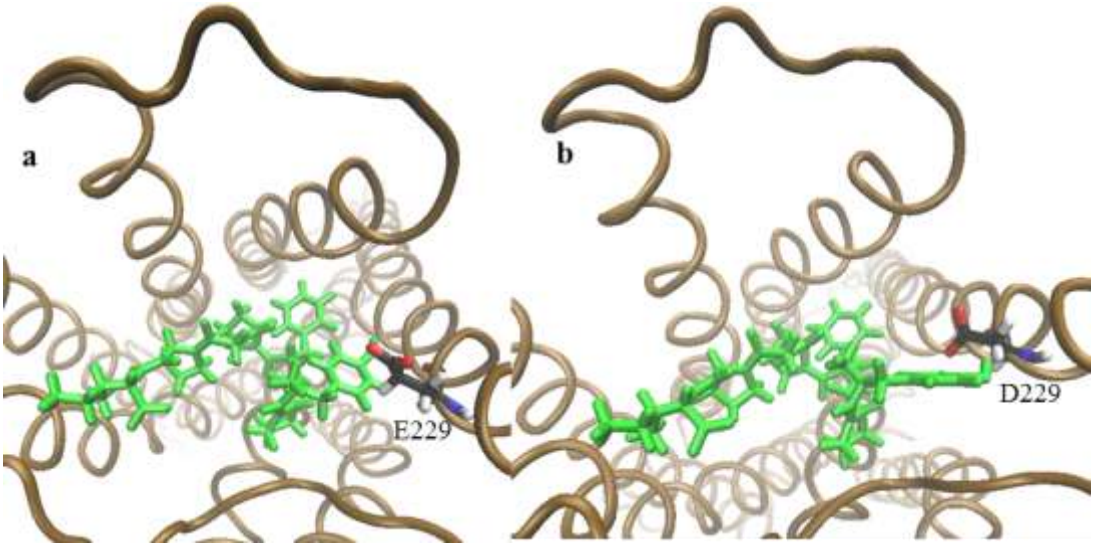



**Figure 79: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor and b) E229D μ receptor.**

In the WT μ receptor, Glu229 makes a hydrophobic interaction with its ethyl side chain and the $C_\beta$ of Tyr of β-casomorphin-7 (Figure 79a). When rescored in the E229D mutant, a steric clash occurred between the side chain of the D229 and the phenolic group of the Tyr of β-casomorphin-7 (Figure 79b).

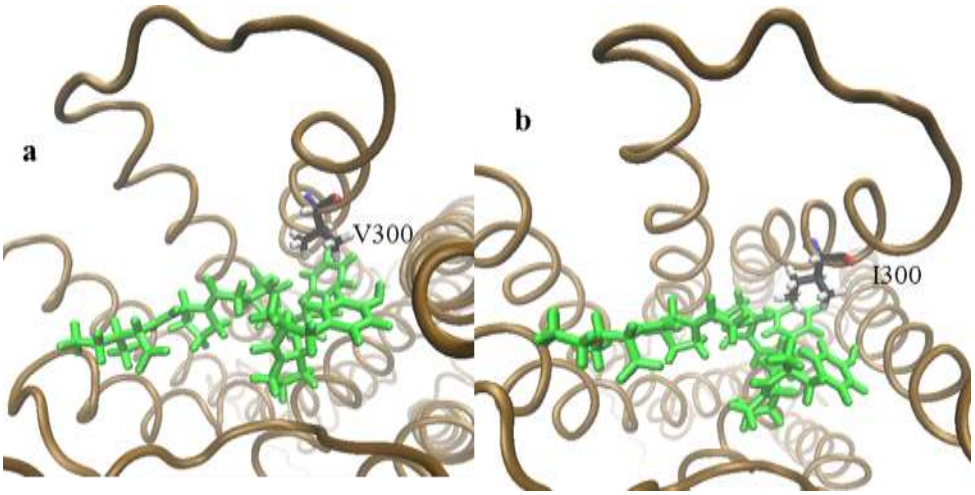



**Figure 80: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor and b) V300I μ receptor.**

In the WT μ receptor, the Val300 side chain makes a hydrophobic interaction with the aromatic ring of the Phe of β-casomorphin-7 (Figure 80a). The substitution of this Val to Ile (V300I) resulted in a steric clash between the Ile sidechain and the aromatic ring of the Phe of β-casomorphin-7 (Figure 80b), resulting in a negative score.

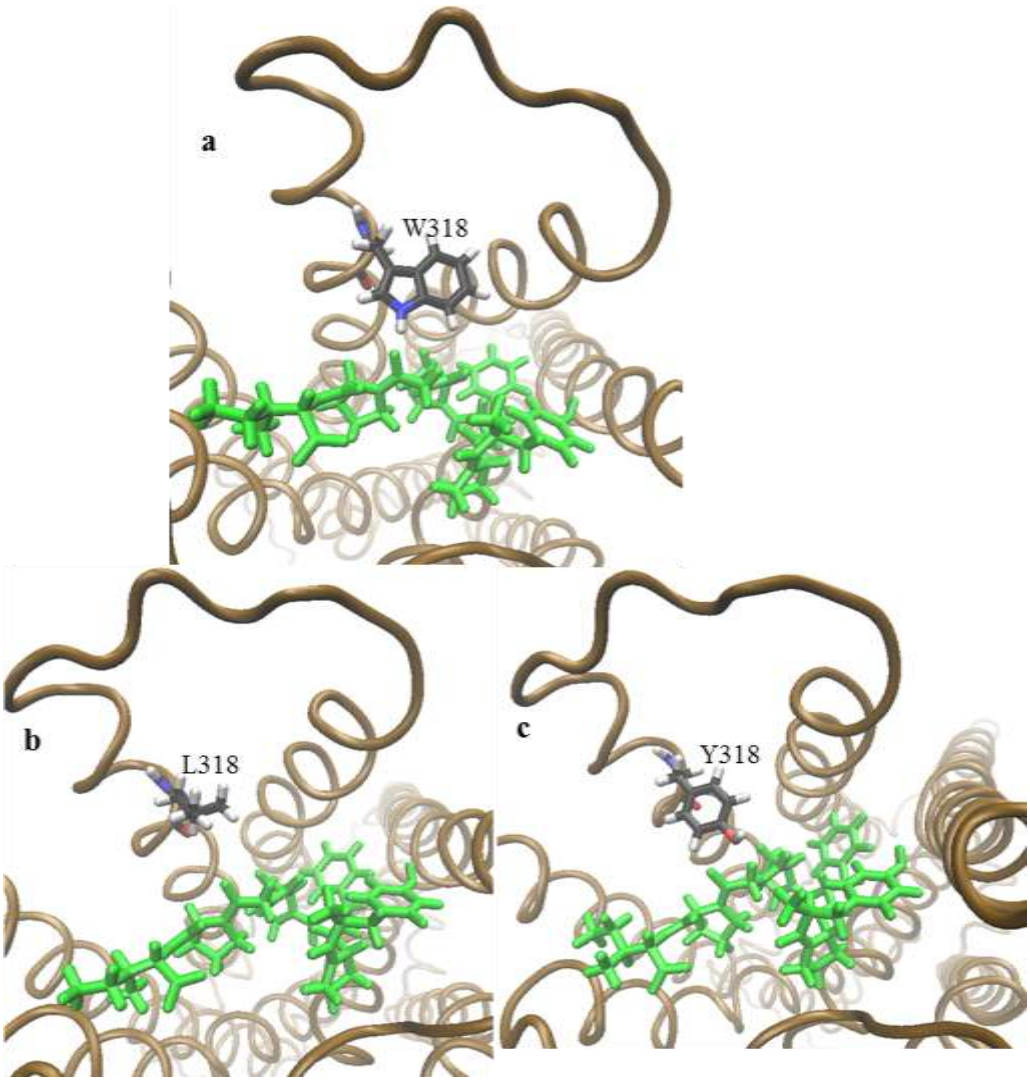



**Figure 81: Visual representation of β-casomorphin-7 (green) pose 235 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor.**

In the WT μ receptor, the Trp318 indole N-H makes an H-bond with the carbonyl group of the Gly of β-casomorphin-7 (Figure 81a). The Leu residue of the W318L mutant makes no interactions as was observed with the δ receptor rescoring (Figure 81b). Likewise, the Tyr of the W318Y mutant was seen clashing with the Gly of β-casomorphin-7 (Figure 81c); however in the κ receptor rescoring, this Tyr residue did not clash with the Gly but rather with the second Pro of β-casomorphin-7 (Figure 64f).

Based on the μ receptor mutant rescores of pose 235, it was concluded that the three amino acid residues Glu229, Val300 and Trp318 are responsible for the discrimination of β-casomorphin-7 between the three receptors for pose 235.

***Pose 12***

Pose 12 of β-casomorphin-7 interacted with two non-conserved amino acids, Asn230 and Trp318, that discriminate between the receptors. Therefore, three mutants of the WT μ receptor were introduced, N230L, W318L and W318Y. For the residue Asn230, only the κ receptor mutant (N230L) was generated, as only the κ receptor showed discrimination at this position.

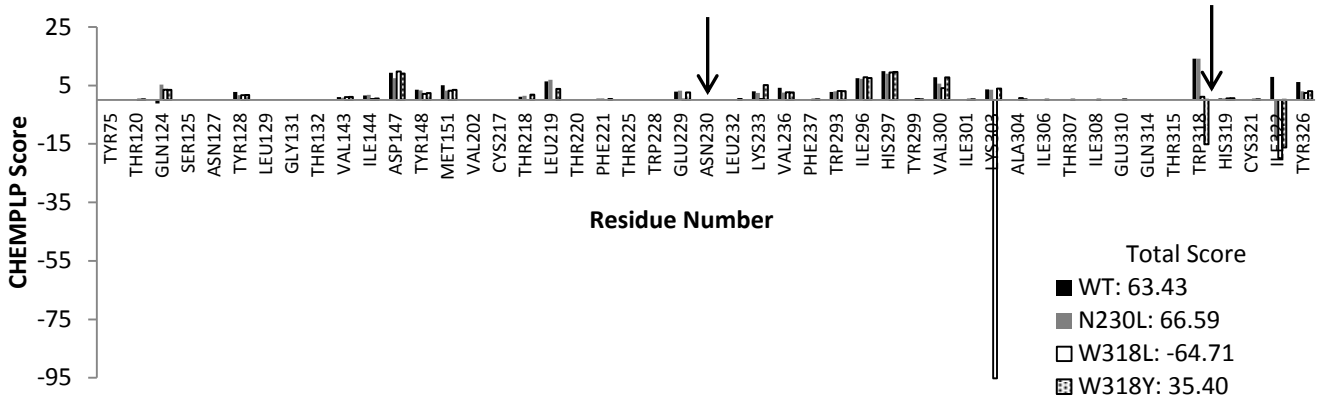



**Figure 82: Residue scores from rescoring of the μ receptor pose 12 in mutant μ receptors.**

Like with pose 9, both the WT μ receptor and the N230L mutant have a score of zero at the position of 230. Position 318 shows a similar score pattern as in the δ and κ receptor rescoring for this pose (Figure 61d and Figure 82), bearing a small positive score for Leu and a negative score for Tyr respectively (Figure 82).

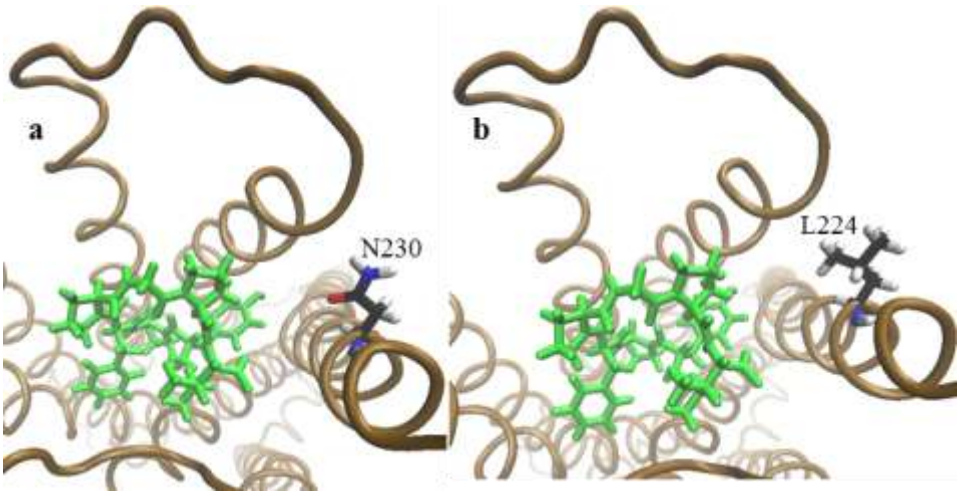



**Figure 83: Visual representation of β-casomorphin-7 (green) pose 12 in the μ receptor rescored in a) WT μ receptor and b) N230L μ receptor.**

Asn230 of the WT μ receptor makes no interaction with β-casomorphin-7, as mentioned previously (Figure 83a). The Leu230 in the N230L mutant does not make any interactions with the β-casomorphin-7 either. This is the same observation that was made for pose 9 regarding this residue.

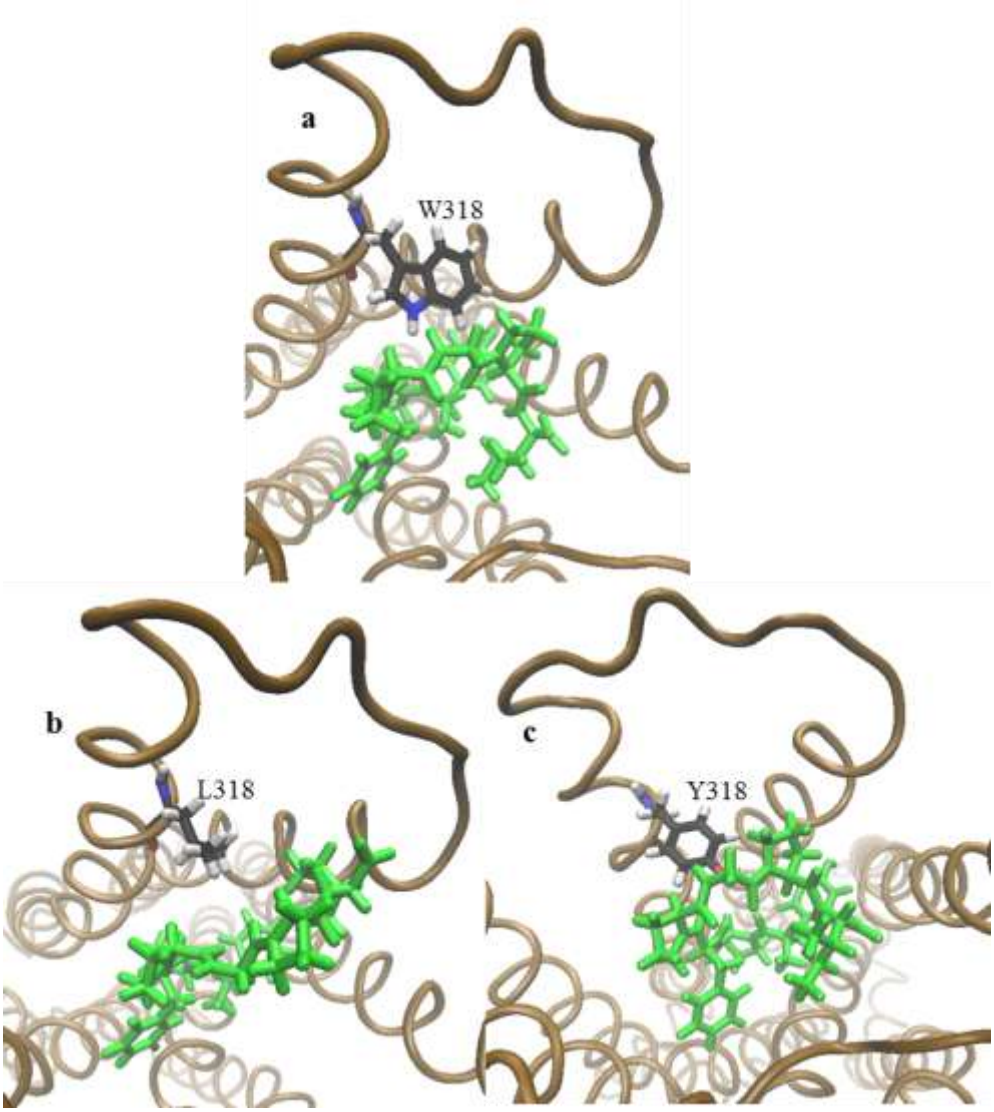



**Figure 84: Visual representation of β-casomorphin-7 (green) pose 12 in the μ receptor rescored in a) WT μ receptor, b) W318L μ receptor, and c) W318Y μ receptor.**

In the WT μ receptor, the indole N-H of the Trp318 forms an H-bond with the carbonyl of the Phe of β-casomorphin-7 (Figure 84a). Mutating this residue to Leu (W318L) resulted in a very small positive score, because of the large distance between the Leu and β-casomorphin-7 (Figure 84b). Mutation of Trp to Tyr (W318Y) resulted in a negative score, which was caused by steric clash between the phenolic OH of the Tyr318 and the Gly of β-casomorphin-7 (Figure 84c). Both of these mutants μ receptor rescores (W318L and W318Y), have similar score

changes to that of the δ and κ receptor rescoring (Figure 61d), compared to that of the WT μ receptor for the residue aligned with Trp318.

Based on the rescoring of pose 12 in these μ receptor mutants, the amino acid residue Trp318 was found to be responsible for the discrimination of β-casomorphin-7 between the three receptors.

**Table 14: List of the amino acids of the four poses, 468, 9, 235 and 12 predicted to the responsible for discriminating β-casomorphin-7 between the three receptors.**

| | Amino Acid Residues Predicted to Discriminate Between Receptors |
|---|---|
| Pose 468 | Lys303, Trp318 |
| Pose 9 | Trp318 |
| Pose 235 | Glu229, Val300, Trp318 |
| Pose 12 | Trp318 |

Trp318 is an amino acid residue shared among the four poses predicted to be responsible for discriminating between the three receptors (Table 14, black arrow in Figure 60).

## Discussion

### 1: Non-Native β-FNA Docking

Each of the crystal structure ligands (Figure 12) is selective for its corresponding receptor. The purpose of performing non-native docking experiments was to verify that GOLD could differentiate the binding affinities of these ligands to each of the receptors based on the docking score. β-FNA binds selectively to the μ receptor. The κ receptor scored lower compared to the μ receptor for the β-FNA ligand as expected, however, the δ receptor scored higher compared to the μ receptor for the β-FNA which was an unexpected result (refer to Results: Section 3b).

The reason the score distribution of the κ receptor was lower than the μ and δ receptor could the easily determined. The crystal structure pose of β-FNA in the μ receptor was sterically impossible to achieve in the κ receptor. This was shown by rescoring the top scoring pose from the μ receptor docking (Figure 30) into the κ receptor (Figure 85a).

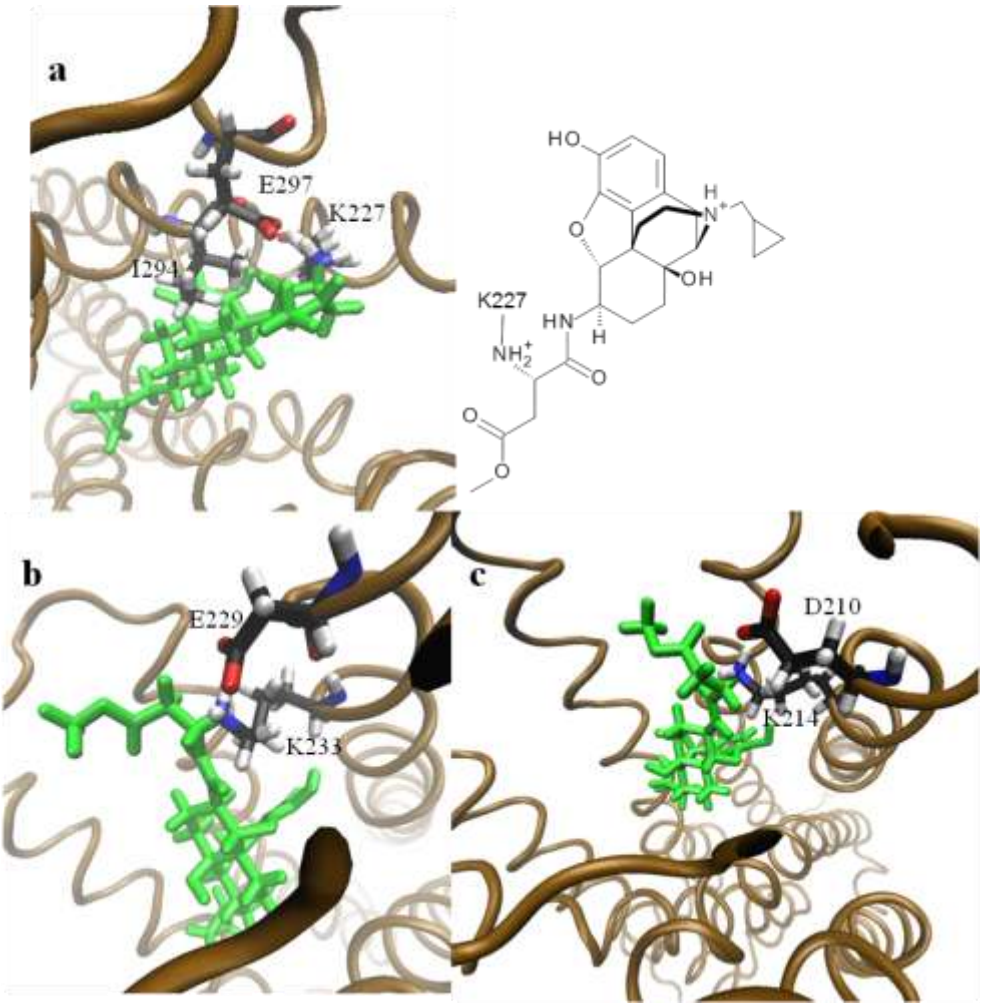



**Figure 85: β-FNA a) Rescoring of top scoring μ receptor pose in the κ receptor, b) Top scoring β-FNA pose in δ receptor, and c) Top scoring β-FNA pose in μ receptor.**

The Ile294 clashes with the ring system of β-FNA. The μ and δ receptors have a Val at this position which is not as large as Ile. Additionally, Glu297 forms a dipole-dipole repulsion with the amide carbonyl oxygen, further disfavouring the pose.

β-FNA scores higher in the δ receptor (average score :146.1) than in the μ receptor (average score: 132.7). This result was unexpected as β-FNA is a ligand that binds selectively to the μ receptor (Figure 37). A plausible explanation for this observation is that residue Asp210, which aligns with Glu229 in the μ receptor, has a shorter side chain, and thus is able to make a salt bridge with the Lys233 side chain ammonium, which in turn covalently links β-FNA to the receptor (Figure 85b). Glu229 of the μ receptor, which has a longer side chain, on the other hand, could not make a strong salt bridge due to the sub-optimal distance to the ammonium centre (Figure 85c). To verify this proposal, the per atom scores were analyzed. However, the δ receptor had a worse score for the ammonium centre than the μ receptor which disproved this theory. It was observed that in the δ receptor, the geometry of the ammonium is correct (tetrahedral) (Figure 85b). In the μ receptor, on the other hand, the geometry was a much distorted tetrahedral structure (Figure 85c). This resulted in a lower score as this geometry is not normal for an ammonium species. The distorted tetrahedral geometry only requires the positions of the hydrogen atoms to be corrected however GOLD cannot make this geometric correction at covalent linkage points. When GOLD generates the covalent bond with this Lys residue, it can rotate the ligand about the linked atom on all three axes while searching for the optimal ligand-receptor interactions. This can result in abnormal geometry at the linking atom. This is one of the limitations of GOLD in covalent docking experiments.

Even with this geometric flaw in the docking poses, this does not explain why β-FNA does not bind to the δ receptor experimentally. The answer to this could lie in the way that β-FNA is linked to the receptor, that is through an α,β-unsaturated ester linkage. The olefinic moiety can undergo a Michael-type addition to the Lys residue. Chen and co-workers showed that even if the Lys is mutated to an Ala, the β-FNA will still bind with high affinity as a

reversible antagonist to the μ receptor[80]. This observation suggested that the core ring system of β-FNA first binds through non-covalent interactions to the binding pocket. Only when this occurs does the covalent linkage form. The reason that the β-FNA does not bind to the δ receptor experimentally may be because of a steric or repulsive clash between the δ receptor and the unreacted β-FNA which is absent in the μ receptor.

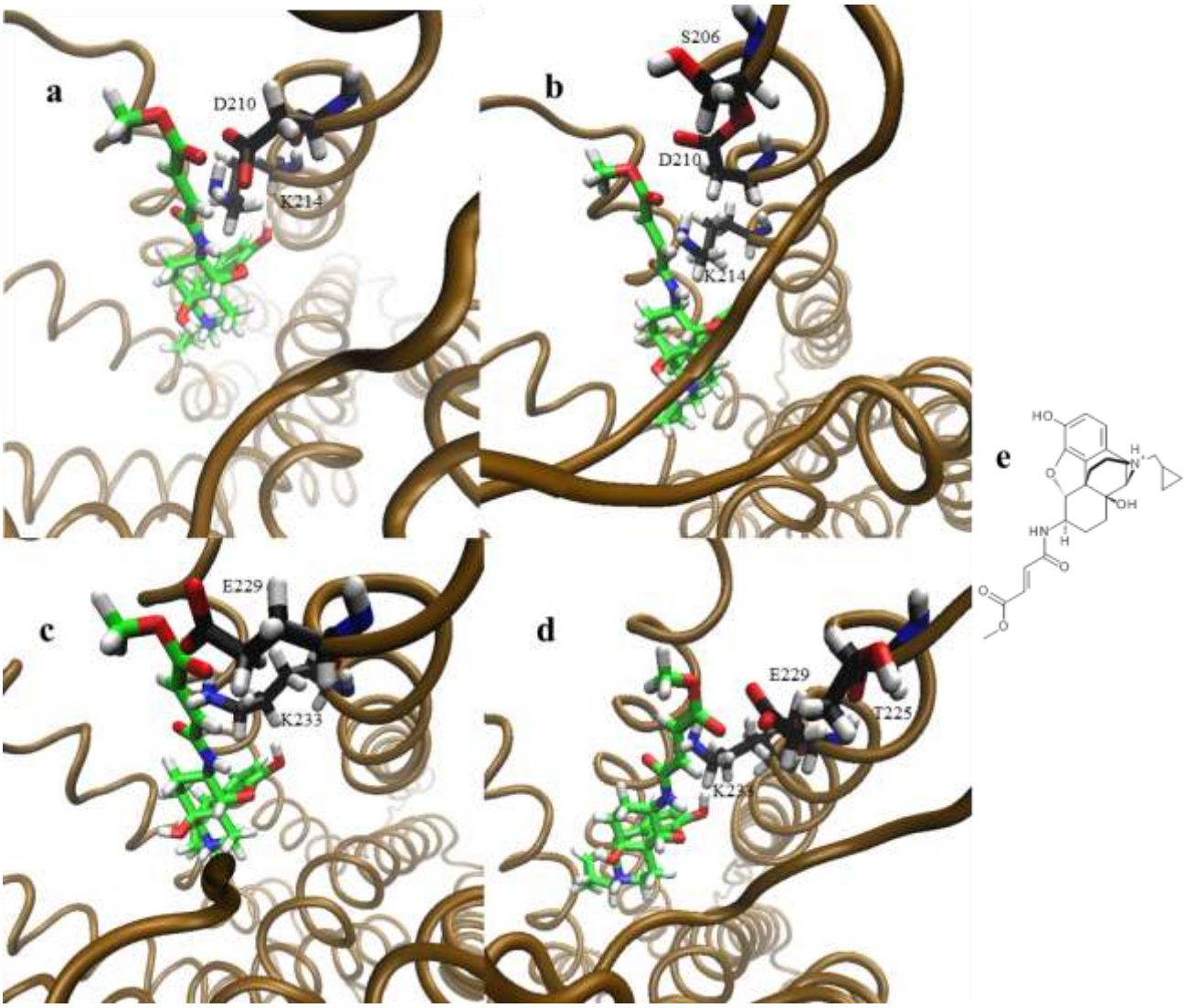



**Figure 86 Comparison of unreacted β-FNA in the μ verses δ receptor. a) Unreacted β-FNA (green scaffold) in the δ receptor, b) unreacted β-FNA (green scaffold) in the δ receptor with an alternative conformation of the side chain of the residue Asp210, c) unreacted β-FNA (green scaffold) in the μ receptor, d) unreacted β-FNA (green scaffold) in the μ receptor with an alternative conformation of the side chain of the residue Glu229, and e) 2-dimensional structure of the unreacted β-FNA.**

Building on this hypothesis, the top scoring pose of β-FNA from each of the two receptors (δ and μ) was taken. The side chain of β-FNA which bears the covalent linkage was deleted and then reconstructed as the α,β-unsaturated ester. This allowed the ring system of β-FNA to remain in the same position in the binding site. This represents the pose that α,β-

unsaturated ester of β-FNA would have to adapt in order to form a covalent bond. In the δ receptor, residue Asp210 becomes problematic in this pose (Figure 86a), as Asp210 forms a dipole-dipole repulsion with the ester group. While the Asp residue can alter its side chain conformation to alleviate this repulsion, such conformational changes only direct Asp210 towards the α-helix (TMV) and cause even worse dipole-dipole repulsion with the backbone oxygen of Ser106 (Figure 86b). In the μ receptor crystal structure, the Glu229 also forms a dipole-dipole repulsion with the ester group (Figure 86c), however the additional methylene group in the side chain of Glu relative to Asp allows for additional flexibility. As such, Glu could alter its conformation such that it would neither conflict with the ester group or the residue Thr225 (at the same position as Ser106 in the δ receptor) (Figure 86d).

This hypothesis was further analysed quantitatively via rescoring. In the crystal structure of the μ receptor (Figure 86c), the Glu229 side chain forms a disfavoured interaction with β-FNA bearing a large negative score. In the alternative conformation (Figure 86d), the score was greatly improved. This conformation was generated by uploading the coordinates of the Glu229 side chain atoms into ArgusLab, which has a tool that can change dihedral angles. Once the dihedral angles were changed, the Glu side chain was saved as coordinates which were subsequently substituted for the original Glu229 side chain coordinates. This side chain was also scored against the rest of the receptor in order to compare the extent of the conflict that the side chain conformation has with other residues in the receptor. Only a minor decrease in the score was observed (Figure 87a). These results suggest that Glu229 is able to adopt a conformation that permits the unreacted β-FNA to bind without causing a major conflict between β-FNA and other residues of the receptor.

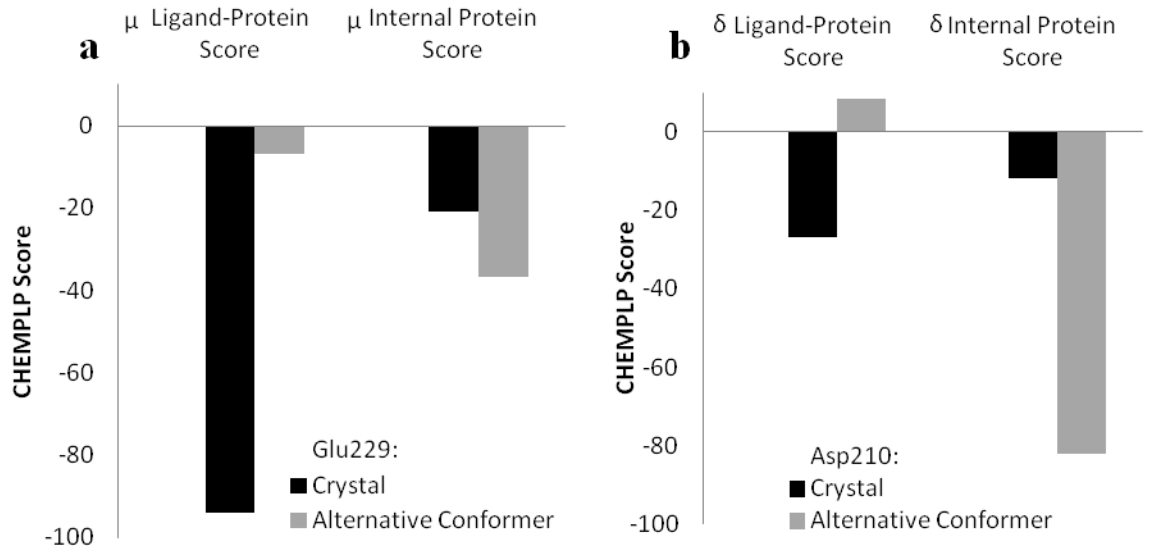



**Figure 87: Quantitative analysis of top scoring β-FNA pose in the a) μ and b) δ receptors. Ligand-Protein Score is the score between the receptor and the β-FNA and the Internal Protein Score is the score between the Glu229 (μ) or Asp210 (δ) side chains and the rest of the receptor.**

The side chain of the residue Asp210 of the δ receptor also forms a similar conflict with β-FNA. Adjusting the conformation of this side chain alleviates this conflict and results in a slightly positive score. This conformation was generated the same way as the Glu229 conformer of the μ receptor. However, when the side chain is scored against the rest of the receptor, the score was greatly decreased indicating that this conformation is disfavoured because of an internal conflict within the protein (Figure 87b). Two other conformations of Asp210 were tested by the same methodology as the alternative conformation of Asp210 and both of them involved in a conflict with another residue of the receptor or β-FNA. It was concluded that none of the conformations of the Asp210 side chain can be accommodated without conflict with the unreacted β-FNA and other residues of the receptor.

It was also determined that the Lys residue only forms the covalent linkage at one of the two stereocentres because only one face of the olefin is accessible to the Lys. When β-FNA is

docked into the receptor, the olefinic carbons must be converted to their tetrahedral forms that are present in the ligand's bound state, as the docking program cannot simulate the formation of this bond. This operation grants more flexibility to the ester group, making it impossible to differentiate between the poses between the μ and δ receptors. Therefore, it was concluded that β-FNA cannot bind to the δ receptor because the olefin bearing side chain could not fit in the binding pocket. The same can be concluded for the κ receptor as it also bears an Asp at the same position as the δ receptor.

## 2: DAMGO Pose Analysis

From previous discussion (refer to Results: Section 5b), two DAMGO poses, 463 and 94, are considered to be consistent with the work by Minami and co-workers[32] determining that Asn127 in the μ receptor (Lys108 in the δ receptor) discriminates DAMGO binding to the μ and δ receptors, and Seki and co-workers[31] who determined that four mutations (E297K, S310V, Y312W and Y313H) on the ECLIII of the κ receptor are necessary to give it similar affinity for DAMGO as the μ receptor. In addition, Surratt and co-workers[55] identified the conserved residues Asp147 and His297 as being important for the binding of DAMGO to the μ receptor. It then becomes necessary to determine which of the two poses is a better representation of the actual pose of DAMGO in the μ receptor. The work by Minami and co-workers also suggested that mutations on the ECLIII of the δ receptor to their μ receptor counterparts had no effect on DAMGO binding of the mutant δ receptor[32]. The only non-conserved amino acid residue of the ECLIII of the μ receptor that has a score in poses 463 and 94 is Trp318. This amino acid residue should have a similar score as the δ receptor counterpart (Leu300). This is true for pose 463 where both the μ and δ receptors had minimal score at this position (Figure 53a). It was not true, however,

for pose 94 where the position of Trp318 of the μ receptor showed a much higher score compared with the Leu300 of the δ receptor (Figure 53b) due to their size differences.

In addition, Val300 is conserved between the μ and δ receptors (Val281 in δ receptor) but had large score differences between the two receptors in pose 94 (Figure 53b). Pose 463, on the other hand, had the same score for the two receptors for this residue (Figure 53a). Therefore, it was concluded that pose 463 was a better representation of the actual pose of DAMGO in the binding pocket of the μ receptor.

### 3: β-Casomorphin-5 Pose 268 in the W318Y Mutant

Rescoring pose 268 of β-casomorphin-5 in the W318Y mutant μ receptor resulted in a small positive score for the residue Tyr318 (Figure 58) compared with the score in the κ receptor which had a considerable negative score (Figure 56c). This finding suggested a major conformational difference in these residues between these two receptors. This difference was verified by superimposing the κ receptor over the W318Y μ receptor mutant using the same criteria as for the WT receptor superposition (refer to Results: Section 1).

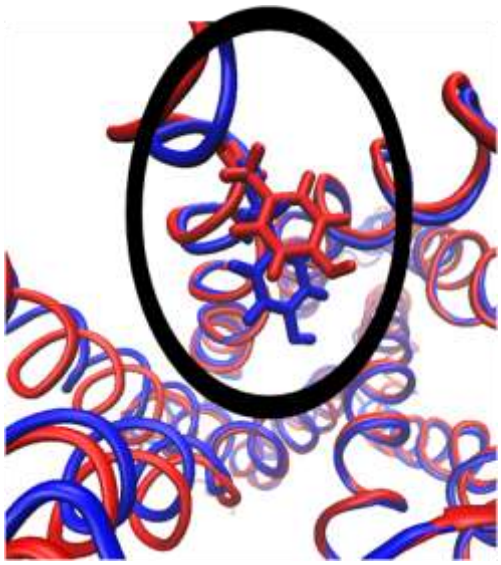


**Figure 88: Superposition of W318Y μ receptor (red) and WT κ receptor (blue). The amino acid residue Tyr318 (μ) and Tyr312 (κ) are shown.**

It can be seen from Figure 88 that the conformation of this Tyr residue is different. The backbone atoms of the local segment (Phe313-His319 for the μ receptor, Ala307-Tyr313 for the κ receptor) of the α-helix (TMVII) were not superimposed (backbone atom RMSD: 2.89Å) even though the central segments (Phe320-Phe338 for the μ receptor, Phe314-Phe332 for the κ receptor) of the α-helix superimposed very well (backbone atom RMSD: 0.64Å) (Figure 88). The fact that the amino acid residue is not superimposed in the two receptors explains the difference in score. Different interactions have formed as a result of different positions of these Tyr side chains with respect to β-casomorphin-5.

## 4: β-Casomorphin-7 Pose Selection

While it was straightforward to determine the best pose to represent the actual β-casomorphin-5 pose in the μ receptor, it is more complicated to select the pose that can best represent β-casomorphin-7. All four poses analysed showed discrimination between the three receptors. Among these, Trp318 was implicated in receptor discrimination for all four poses.

The four poses were compared with the predicted pose of β-casomorphin-5. Due to the sequence similarity of these two peptide ligands, β-casomorphin-5 and 7 are expected to have some pharmacophoric similarity. Four components of the two casomorphins are compared: N-terminus, C-terminus, the phenolic group of Tyr and the phenyl ring of Phe.

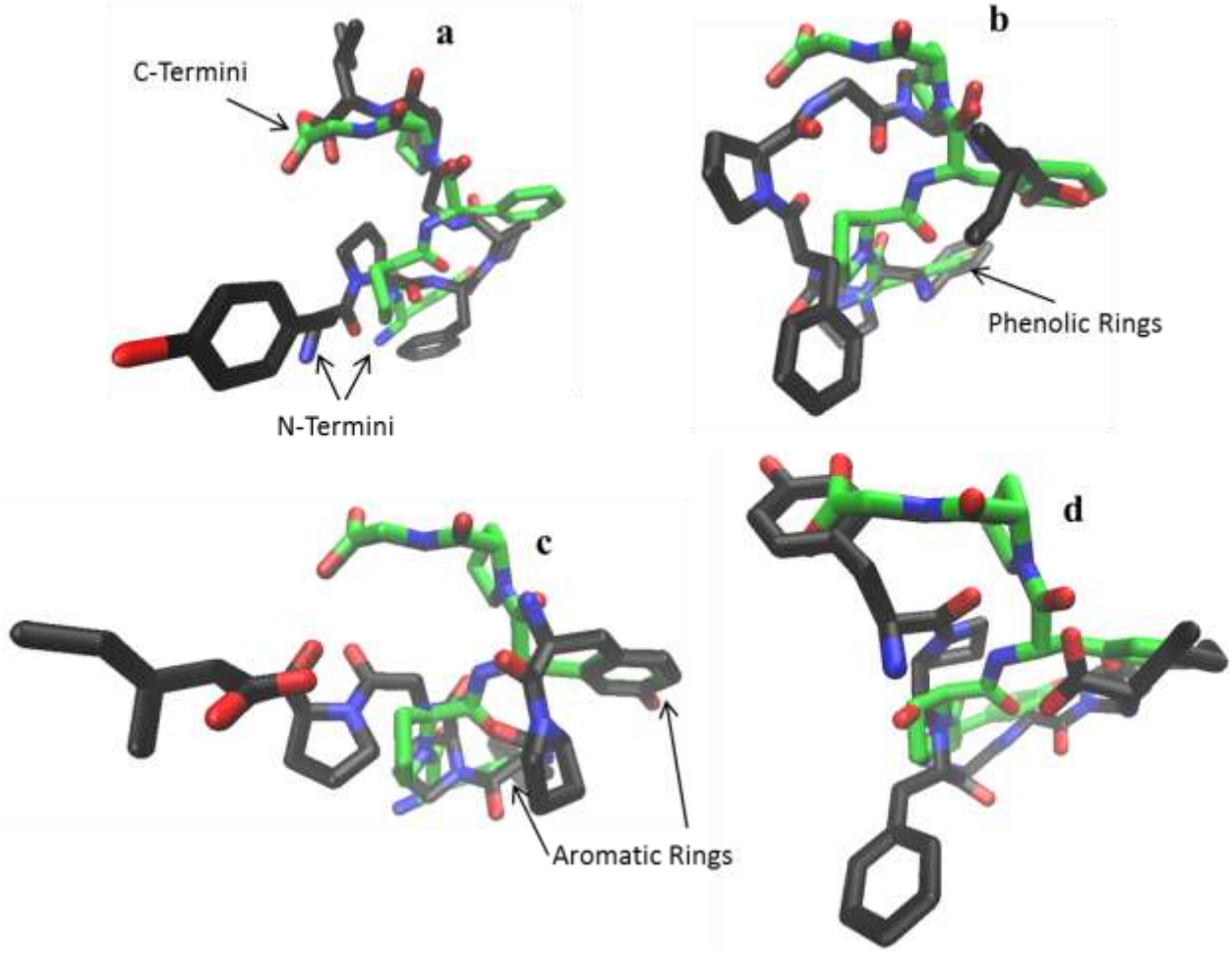



**Figure 89: Superposition of pose 268 of β-casomorphin-5 (green scaffold) with a) pose 468, b) pose 12, c) pose 235, and d) pose 9 of β-casomorphin-7 (black scaffolds).**

Pose 468 has no similarity with pose 268 of β-casomorphin-5 at the N-terminus. At the C-terminus, however there was significant overlap of the backbone. The C-termini themselves were superimposed along with their preceding Pro residue (Table 15). In addition, the second Pro

residue of β-casomorphin-7 superimposes well with the phenyl ring of Phe of β-casomorphin-5. The first Pro residues of the two ligands also superimposed well. While the N-termini faced in the same general direction, they interacted with different residues in the binding pocket (Figure 89a). Poses 12 and 268 only superimpose at the phenolic rings (Table 15). There was no other similarity between pose 12 of β-casomorphin-7 and pose 268 of β-casomorphin-5 (Figure 89b). In pose 235, only the two aromatic rings superimpose well, however they did so in an opposite position, that is, the position that had a phenolic group for one of the ligands had the phenyl group for the other ligand (Figure 89c). For pose 9 and 268, no superposition of the molecules was observed. The only commonality between the poses of the two ligands was that β-casomorphin-5 had its phenyl ring superimposed onto the Ile side chain of β-casomorphin-7. Both of these amino acids side chains are hydrophobic groups of similar size (Figure 89d). Due to the lack of similarity between pose 268 of β-casomorphin-5 and poses 12, 235 and 9 of β-casomorphin-7, these three β-casomorphin-7 poses are possibly not good representations of the actual pose of β-casomorphin-7 in the μ receptor even though they did show discrimination between the receptors. Pose 468, however, had similarities to several sections of β-casomorphin-5, which makes this pose a good tentative proposal for the actual pose of β-casomorphin-7 in the μ receptor.

**Table 15: The comparison of the RMSD of various components of β-casomorphin-5 and 7. Row 3 compares the backbone atom (N, $C_\alpha$, C and O) of the Phe-Pro-Gly sequence of β-casomorphin-5 and the Gly-Pro-Ile sequence of β-casomorphin-7.**

| Atoms Included in RSMD Calculation | RMSD (Å) | | | |
|---|---|---|---|---|
| | Pose 468 | Pose 9 | Pose 235 | Pose 12 |
| Tyr backbone | 4.87 | 7.00 | 8.31 | 2.25 |
| Tyr side chain | 13.21 | 12.08 | 8.29 | 1.92 |
| Backbone atoms of the last three amino acids of the C-terminus | 1.34 | 7.41 | 8.81 | 6.29 |

The Tyr residues of pose 268 of β-casomorphin-5 and pose 468 of β-casomorphin-7 face opposite directions. Yet, it was shown in experimental studies that this Tyr residue is essential for binding, where no measurable binding is seen in its absence[82]. In these two poses the side chains of the Tyr residue make contact with the edges of the binding pocket. This observation would suggest that these Tyr residues functioned as a brace for the rest of the ligand and therefore do not necessarily make a conserved interaction.

Neither the proposed β-casomorphin-5 nor 7 pose has their first Pro residue in cis conformation contrary to what was proposed by Borics and Tóth[81]. An explanation could be that Borics and Tóth did not use a receptor in their study whose presence can influence the stability cis tertiary amide bond.

## 5: Global Receptor Conformation

It is well known that proteins have multiple conformations. Changes in the conformations often involve conformational changes in the backbone of the proteins. GOLD is unable to make changes to the backbone structure. Therefore it was assumed that the conformational differences in the backbone structure were negligible between the active versus inactive form of the receptors. In order to determine whether this assumption is plausible or not, morphine was docked in the μ receptor. Morphine is an alkaloid that is structurally similar to that of naltrindole and β-FNA (Figure 90) but is an agonist, unlike the other two alkaloids. It has no rotatable bonds (except the OH groups) so it should give a very consistent pose unlike what was seen for β-casomorphins. The same charged state of the ammonium centre was used as for the antagonist alkaloids.

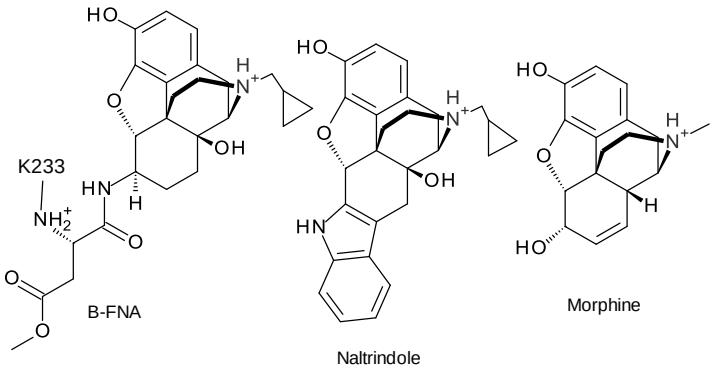



**Figure 90: Structures of alkaloid ligands used.**

Clusters of three different poses were obtained. The clusters were determined by displaying all the docked poses of morphine simultaneously, revealing three geometrically distinct sets of poses by visual inspection. To determine the quantity of poses in each of the clusters, an RMSD calculation was carried out. In this scenario, there is no crystal structure pose to use as a reference, therefore a random docked pose was selected and used as the reference pose for the RMSD calculation. As such, the RMSD value was not as important as the distribution pattern (Figure 91).

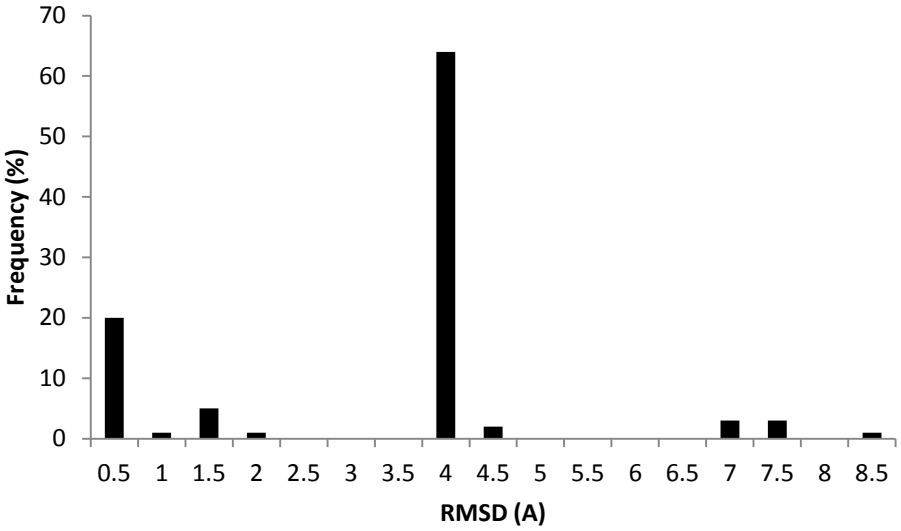



**Figure 91: Distribution of the three poses of morphine in the μ receptor.**

These results from Figure 91 can be compared with the same style of graph generated from the naltrindole docking to the δ receptor (Figure 26). Protonated naltrindole adapts a single pose 90% of the time in the δ receptor. Morphine will only adapt a single pose 66% of the time in the μ receptor. Furthermore, none of these poses is similar to that of β-FNA nor is any of them in the same section of the binding pocket as β-FNA is covalently bound. The other stereoisomer of the ammonium centre was also docked but gave many more poses and thus was not investigated further.

In order to compare the results from the μ receptor with that of the δ and κ receptors, morphine was docked in the δ and κ receptors and their CHEMPLP score distributions were compared (Figure 92). The κ receptor docking resulted in four different poses located in two different locations within the binding pocket with a very disperse distribution of scores. However, the δ receptor only gives a single pose, where the core ring system is superimposed with that of the native ligand naltrindole (Figure 93).

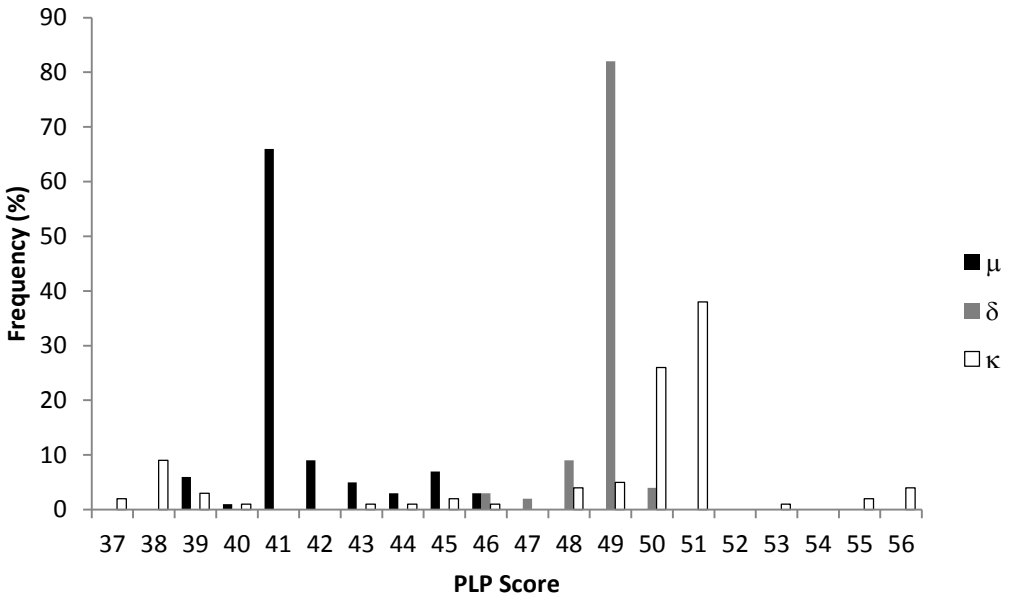



**Figure 92: Comparison of the CHEMPLP score distribution of 100 dockings of morphine in the μ, δ and κ receptors.**

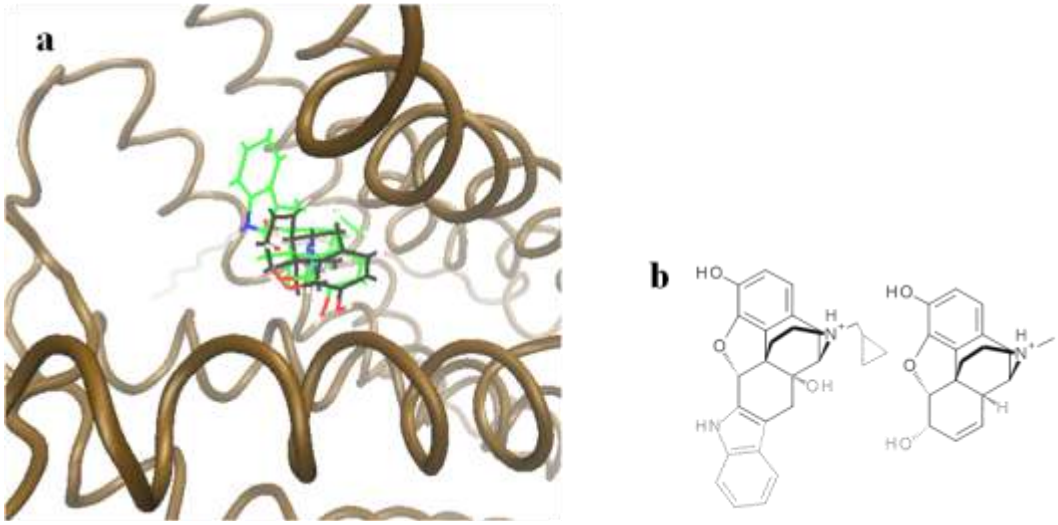



**Figure 93: Superpositioning of naltrindole and morphine a) Comparison of the docked pose of morphine (black scaffold) in the δ receptor and the crystal structure pose of naltrindole (green scaffold), RMSD=0.58Å and b) Atoms used for an RMSD calculation (black), all hydrogen atoms were excluded from the calculation.**

As can be seen from Figure 92, the μ receptor had the lowest score for the morphine docking and it could not be docked in a consistent pose. This result is contradictory to what is known experimentally about morphine, as it is highly selective for the μ receptor[82]. This result

strongly suggests that the receptor may be positioned in a wrong conformation for morphine binding, which implies that the conformation could also be incorrect for the β-casomorphin binding since they are agonists as well. These assumptions can only be verified if crystal structures of the μ receptor with morphine and a β-casomorphin are available.

## Conclusion

GOLD was shown to be reliable in predicting the pose of rigid ligands such as the bound ligand of the opioid receptor found in crystal structures (μ receptor: 4DKL, κ receptor: 4DJH, δ receptor: 4EJ4), however, it was found to be unreliable in predicting poses of highly flexible ligands such as DAMGO and the β-casomorphins. This shortcoming was overcome by scoring the pose on a per amino acids basis. This methodology allowed the poses of DAMGO, β-casomorphin-5 and β-casomorphin-7 to be predicted in the μ receptor. DAMGO in the predicted pose makes all interactions that are consistent with literature findings derived from mutational/chimeric experiments. The predicted β-casomorphin-5 and 7 poses both included the same two amino acids residues of the μ receptor, Lys303 and Trp318, that are involved in discriminating among the three receptors. These two residues can be mutated to their δ and κ counterparts in a laboratory experiment to validate this prediction. Future work will involve docking unnatural β-casomorphins into the μ receptor, such as D-Pro$^2$ β-casomorphin-5, D-Phe$^3$ β-casomorphin-5, β-casomorphin-4-amide as well as cyclic forms of these β-casomorphin analogues. These analogues have shown to have higher binding affinity to the μ receptor compared to the unmodified β-casomorphins[82]. These docking experiments would aim to determine why these analogues have higher binding affinity to the μ receptor. The methodology that was used only analyzed a few poses out of the 500 poses generated by GOLD. An RMSD

calculation can be made between each of the analysed poses and whole set of 500 poses for DAMGO, β-casomorphin-5 and β-casomorphin-7 in order to see if are any poses that are geometrically similar. The longer term objective would be to develop an automated system to perform this per amino acid scoring analysis rather than the manual method that was employed here. Another improvement that can be made to the methodology is to allow flexible side chains when rescoring a ligand in a different receptor, which is not permitted in GOLD.

## References


1. Davis, M. P. Opioid Receptor Targeting Ligands for Pain Management: A Review and Update. *Expert Opinion on Drug Discovery* **2010**, *5*, 1007-1022.

2. Klockgether-Radke, A. P. F. W. Sertürner and the Discovery of Morphine; 200 Years of Pain Therapy with Opioids. *Anästhesiol Intensivmed Notfallmed Schmerzther* **2002**, *37*, 244-249.

3. Praag, D. V.; Simon, E. J. Studies on the Intracellular Distribution and Tissue Binding of Dihydromorphine-7,8-$H^3$ in the Rat. *Experimental Biology and Medicine* **1966**, *122*, 6-11.

4. Terenius, L. Specific Uptake of Narcotic Analgesics by Subcellular Fractions of the Guinea-Pig Ileum. *Acta Pharmacologica et Toxicologica* **1972**, *31*, 50-50.

5. Terenius, L. Stereospecific Uptake of Narcotic Analgesics by a Subcellular Fraction of the Guinea-Pig Ileum. *Upsala Journal of Medical Sciences* **1973**, *78*, 150-152.

6. Goldstein, A.; Lowney, L. I.; Pal, B. K. Stereospecific and Nonspecific Interactions of the Morphine Congener Levorphanol in Subcellular Fractions of Mouse Brain. *Proceedings of the National Academy of Sciences of the United States of America* **1971**, *68*, 1742-1747.

7. Martin, W. R.; Eades, C. G.; Thompson, J. A.; Huppler, R. E.; Gilbert, P. E. The Effects of Morphine- and Nalorphine- Like Drugs in the Nondependent and Morphine-Dependent Chronic Spinal Dog. *The Journal of Pharmacology and Experimental Therapeutics* **1976**, *197*, 517-532.

8. Lord, J. A. H.; Waterfield, A. A.; Hughes, J.; Kosterlitz, H. W. Endogenous Opioid Peptide: Multiple Agonists and Receptors. *Nature* **1977**, *267*, 495-499.

9. Hughes, J.; Smith, T. W.; Kosterlitz, H. W.; Hergill, L. A.; Morgan, B. A.; Morris, H. R. Identification of Two Related Pentapeptides from the Brain with Potent Opiate Agonist Activity. *Nature* **1975**, *258*, 577-579.

10. Bauer, U.; Nakazi, M.; Kathmann, M.; Göthert, M.; Schlicker, E. The Stereoselective Kappa-Opioid Receptor Antagonist Mr 2266 does not Exhibit Stereoselectivity as an Antagonist at the Orphan Opioid (ORL1) Receptor. *Naunyn-Schmiedeberg's Archives of Pharmacology* **1999**, *359*, 17-20.

11. Calogero, A. E.; Scaccianoce, S.; Burrello, N.; Nicolai, R.; Muscolo, L. A.; Kling, M. A.; Angelucci, L.; D'Agata, R. The Kappa-Opioid Receptor Agonist MR-2034 Stimulates the Rat Hypothalamic-Pituitary-Adrenal Axis: Studies in vivo and in vitro. *Journal of Neuroendocrinology* **1996**, *8*, 579-585.

12. Kosterlitz, H. W.; Lyndon, R. J.; Watt, A. J. The Effect of Adrenaline, Noradrenaline, and Isoprenaline on Inhibitory alpha- and beta- Adenoreceptors in the Logitudinal Muscle of the Guinea-Pig Ileum. *British Journal of Pharmacology* **1970**, *39*, 398-413.

13. Bunzow, J. R.; Saez, C.; Mortrud, M.; Bouvier, C.; Williams, J. T.; Low, M.; Grandy, D. K. Molecular Cloning and Tissue Distribution of a Putative Member of the Rat Opioid Receptor Gene Family that is not a Mu, Delta or Kappa Opioid Receptor Type. *Federation of European Biochemical Societies* **1994**, *347*, 284-288.

14. Lambert, D. The Nociceptin/Orphanin FQ Receptor: a Target with Broad Therapeutic Potential. *Nature Reviews. Drug Discovery* **2008**, *7*, 694-710.

15. Hirao, A.; Imai, A.; Sugie, Y.; Yamada, Y.; Hayashi, S.; Toide, K. Pharmacological Characterization of the Newly Synthesized Nociceptin/Orphanin FQ–Receptor Agonist 1-[1-(1-Methylcyclooctyl)-4-piperidinyl]-2-[(3R)-3-piperidinyl]-1H-benzimidazole as an Anxiolytic Agent. *Journal of Pharmacological Sciences* **2008**, *106*, 361-368.

16. Hargrave, P. A.; Fond, S.; McDowell, J. H.; Mas, M. T.; Curtis, D. R.; Wang, J. K. The Partial Primary Structure of Bovine Rhodopsin and its Topography in the Retinal Rod Cells Disc Membrane. *Neurochemistry* **1980**, *1*, 231-244.

17. Ovchinnikov, Y. A. Rhodopsin and Bacteriorhodopsin: Structure-Function Relationships *Federation of European Biochemical Societies Letters* **1982**, *148*, 179-191.

18. Bera, I.; Laskar, A.; Ghoshal, N. Exploring the Structure of Opioid Receptors with Homology Modeling Based on Single and Multiple Templates and Subsequent Docking: A Comparative Study. *Journal of Molecular Modeling* **2010**, *17*, 1207-1221.

19. Graham, R. M.; Sena, L.; Schwarz, K. R.; Homey, C. J. Evidence from Photoaffinity-Labeling Studies for Coupling of the Alpha-1-Adrenergic Receptor to a Guanine-Nucleotide (G) Binding-Protein. *Federation Proceedings* **1986**, *45*, 1571-1571.

20. Costa, T.; Klinz, F.; Vachon, L.; Herz, A. Opioid Receptors are Coupled Tightly to G Proteins but Loosely to Adenylate Cyclase in NG108-15 Cell Membranes. *Molecular Pharmacology* **1988**, *34*, 744-754.

21. Allgaier, C.; Daschmann, B.; Sieverling, J.; Hertting, G. Presynaptic K-Opioid Receptors on Noradrenergic Nerve Terminals Couple to G Proteins and Interact with the a2-Adrenoceptors. *Journal of Neurochemistry* **1989**, *53*, 1629-1635.

22. Baldwin, J. M. The Probable Arrangement of the Helices in G Protein–Coupled Receptors. *European Molecular Biology Organization Journal* **1993**, *12*, 1693-1703.

23. Palczewski K; Kumasaka T; Hori T; Behnke C.A.; Motoshima H; Fox; B.A. Crystal Structure of Rhodopsin: A G Protein Coupled Receptor. *Science* **2000**, *289*, 739-745.

24. Dhawan, B. N.; Cesselin, F.; Raghubir, R.; Reisine, T.; Bradley, P. B.; Portoghese, P. S.; Hamon, M. International Union of Pharmacology. XII. Classification of Opioid Receptors. *Pharmacological Reviews* **1996**, *48*, 568-586.

25. McDonald, J.; Lambert, D. Opioid Mechanisms and Opioid Drugs. *Anaesthesia & Intensive Care Medicine* **2008**, *9*, 33-37.

26. Tuteja, N. Signaling Through G Protein Couples Receptors. *Plant Signaling & Behavior* **2009**, *4*, 942-947.

27. Lefkowitz, R. J.; Stadel, J. M.; Caron, M. G. Adenylate Cyclase-Coupled Beta-Adrenergic Receptors: Structure and Mechanisms of Activation and Desensitization. *Annual Review of Biochemistry* **1983**, *52*, 159-186.

28. Willoughby, D.; Cooper, D. M. F. Organization and $Ca^{2+}$ Regulation of Adenylyl Cyclases in cAMP Microdomains. *Physiological Reviews* **2007**, *87*, 965-1010.

29. Rasmussen, S. G. F.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T. A.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K. Crystal Structure of the β2 Adrenergic Receptor-Gs Protein Complex. *Nature* **2012**, *477*, 549-555.

30. Granier, S.; Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Weis, W. I.; Kobilka, B. K. Structure of the δ-Opioid Receptor Bound to Naltrindole. *Nature* **2012**, *485*, 400-404.

31. Seki, T.; Minami, M.; Nakagawa, T.; Ienaga, Y.; Morisada, A. DAMGO Recognizes Four Residues in the Third Extracellular Loop to Discriminate Between Mu- and Kappa-Opioid Receptors. *European Journal of Pharmacology* **1998**, *350*, 301-310.

32. Minami, M.; Nakagawa, T.; Seki, T.; Onogi, T.; Aoki, Y.; Katao, Y.; Katsumata, S.; Satoh, M. A Single Residue, Lys108, of the Delta-Opioid Receptor Prevents the Mu-Opioid Selective Ligand [D-Ala2, *N*-MePhe4, Gly-ol5] Enkephalin from Binding to the Delta-Opioid Receptor. *Molecular Pharmocology* **1996**, *50*, 1413-1422.

33. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega . *Molecular Systems Biology* **2011**, *7*, 1-6.

34. https://www.ebi.ac.uk/Tools/msa/clustalo/.

35. McDonald, J.; Lambert, D. Opioid Receptors. *Continuing Education in Anaesthesia, Critical Care & Pain* **2005**, *5*, 22-25.

36. Davis, M. P. Opioid Receptor Targeting Ligands for Pain Management: A Review and Update. *Expert Opinion on Drug Discovery* **2010**, *5*, 1007-1022.

37. Drenth, J. *Principles of Protein X-Ray Crystallography;* Springer: 2007; .

38. Creighton, T. E. *Proteins: Structures and Molecular Properties;* W. H. Freeman and Co.: 1993

39. Ogata, H.; Nishikawa, K.; Lubitz, W. Hydrogens Detected by Subatomic Resolution Protein Crystallography in a [NiFe] Hydrogenase. *Nature* **2015**, *0*, 1-15.

40. Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: An Environment for Comparative Protein Modeling. *Electrophoresis* **1997**, *18*, 2714-2723.

41. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G, Bhat, T.N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Research* **2000**, *28*, 235-242.

42. Rhodes, G. *Crystallography Made Crystal Clear;* Elsevier Inc.: 2006

43. Thompson, M. A. Molecular Docking Using ArgusLab, an Efficient Shape-Based Search Algorithm and the AScore Scoring Function. *American Chemical Society Meeting* **2004**, *172*, 42.

44. Manglik, A.; Kruse, A. C.; Kobilka, T. S.; Thian, F. S.; Mathiesen, J. M.; Sunahara, R. K.; Pardo, L.; Weis, W. I.; Kobilka, B. K.; Granier, S. Crystal Structure of the μ-Opioid Receptor Bound to a Morphinan Antagonist. *Nature* **2012**, *485*, 321-326.

45. Thompson, A. A.; Liu, W.; Chun, E.; Katritch, V.; Wu, H.; Vardy, E. I.; Huang, X.; Trapella, C.; Guerrini, R.; Calo, G.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Structure of the Nociceptin/Orphanin FQ Receptor in Complex with a Peptide Mimetic. *Nature* **2012**, *485*, 395-399.

46. Wu, H.; Wacker, D.; Mileni, M.; Katritch, V.; Han, G. W.; Vardy, E.; Liu, W.; Thompson, A. A.; Huang, X. P.; Carroll, F. I.; Mascarella, S. W.; Westkaemper. R.B.; Mosier, P. D.; Roth, B. L.; Cherezov, V.; Stevens, R. C. Structure of the Human κ-Opioid Receptor in Complex with JDTic. *Nature* **2012**, *485*, 327-332.

47. Rosenbaum, D. M.; Cherezov, V.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H. J.; Yao, X. J.; Weis, W. I.; Stevens, R. C.; Kobilka, B. K. GPCR Engineering Yields High-Resolution Structural Insights into Beta(2)-Adrenergic Receptor Function. *Science* **2007**, *318*, 1266-1273.

48. Koneru, A.; Satyanarayana, S.; Rizwan, S. Endogenous Opioids: Their Physiological Role and Receptors. *Global Journal of Pharmacology* **2009**, *3*, 149-153.

49. Ananthan, S.; Khare, N. K.; Saini, S. K.; Seitz, L. E.; Bartlett, J. L.; Davis, P.; Dersch, C. M.; Porreca, F.; Rothman, R. B.; Bilsky, E. J. Identification of Opioid Ligands Possessing Mixed mu Agonist/delta Antagonist Activity among Pyridomorphinans Derived from Naloxone, Oxymorphone, and Hydromorphone. *Journal of Medicinal Chemistry* **2004**, *47*, 1400-1412.

50. Hruby, V. J.; Agnes, R. S. Conformation–Activity Relationships of Opioid Peptides with Selective Activities at Opioid Receptors. *Biopolymers (Peptide Science)* **1999**, *51*, 391-410.

51. Lovering, A. L.; Safadi, S. S.; Strynadka, N. C. J. Structural Perspective of Peptidoglycan Biosynthesis and Assembly. *Annual Review of Biochemistry* **2012**, *81*, 451-478.

52. Chen, Y.; Mestek, A.; Liu, J.; Hurley, J. A.; Yu, L. Molecular Cloning and Functional Expression of a m-Opioid Receptor from Rat Brain. *Molecular Pharmacology* **1993**, *44*, 8-12.

53. Evans, C. J.; Keith Jr., D. E.; Morrison, H.; Magendzo, K.; Edwards, R. H. Opioid Receptor by Functional Expression. *Science* **1992**, *258*, 1952-1955.

54. Meng, F.; Xie, G. X.; Thompson, R. C.; Mansour, A.; Goldstein, A.; Watson, S. J.; Akil, H. Cloning and Pharmacological Characterization of a Rat Opioid Receptor. *Proceedings of the National Academy of Sciences* **1993**, *90*, 9954-9958.

55. Surratt, C. K.; Johnson, P. S.; Moriwaki, A.; Seidleck, B. K.; Blaschak, C. J.; Wang; J.B.; Uhl, G. R. Charged Transmembrane Domain Amino Acids are Critical for Agonist Recognition and Intrinsic Activity. *The Journal of Biological Chemistry* **1994**, *269*, 20548-20553.

56. Kyoto University Bioinformatics Center GenomeNet.

57. Brantl, V.; Teschemacher, H.; Blasig, J.; Henschen, A.; Lottspeich, F. Opioid Activities of B-Casomorphins. *Life Sciences* **1981**, *28*, 1903-1909.

58. Brantl, V.; Pfeiffer, A.; Herz, A.; Henschen, A.; Lottspeich, F. Antinociceptive Potencies of beta-Casomorphin Analogs as Compared to Their Affinities Towards Mu and Delta Opiate Receptor Sites in Brain and Periphery. *Peptides* **1982**, *3*, 793-797.

59. Koch, G.; Wiedemann, K.; Teschemacher, H. Opioid Activities of Human beta-Casomorphins. *Naunyn-Schmiedeberg's Archives of Pharmacology* **1985**, *331*, 351-354.

60. Liebmann, C.; Szucs, M.; Neubert, K. Opiate Receptor Binding Affinities of Some D-Amino Acid Substituted B-Casomorphin Analogs. *Peptides* **1986**, *7*, 195-199.

61. Ruthrich, H.; Grecksch, G.; Schmidt, R.; Neubert, K. Linear and Cyclic B-Casomorphin Analogues with High Analgesic Activity. *Peptides* **1992**, *13*, 483-485.

62. Leach, A. R. Ligand Docking to Proteins with Discrete Side-Chain Flexibility. *Journal of Molecular Biology* **1994**, *235*, 345-356.

63. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J. A Geometric Approach to Macromolecule-Ligand Interactions. *Journal of Molecular Biology* **1982**, *161*, 269-288.

64. Leach, A. R.; Kuntz, I. D. Conformational-Analysis of Flexible Ligands in Macromolecular Receptor-Sites. *Journal of Computational Chemistry* **1992**, *13*, 730-748.

65. Oshiro, C. M.; Kuntz, I. D.; Dixon, J. S. Flexible Ligand Docking Using a Genetic Algorithm. *Journal of Computer-Aided Molecular Design* **1995**, *9*, 113-130.

66. Jones, G.; Willett, P.; Glen, R. C. Molecular Recognition of Receptor Sites using a Genetic Algorithm with a Description of Desolvation. *Journal of Molecular Biology* **1995**, *245*, 43-53.

67. Zhao, Y.; Sanner, M. F. FLIPDock: Docking Flexible Ligands into Flexible Receptors. *Proteins: Structure, Function, and Bioinformatics* **2007**, *68*, 726-737.

68. Korb, O.; Stutzle, T.; Exner, T. E. Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS. *Journal of Chemical Information and Modeling* **2009**, *49*, 84-96.

69. Blaney, J. M.; Critchlow Jr., R. E.; Judson, R.; Larter, R.; Martin, E. J.; Showalter, K.; Smith, S. J.; Spellmeyer, D. C.; Sutcliffe, B. T.; Topper, R. Q. *Reviews in Computational Chemistry;* VCH: New York, 1997

70. The Official UCSF DOCK Web-site. *http://dock.compbio.ucsf.edu/*.

71. FlexX: Molecular Docking. *http://www.zbh.uni-hamburg.de/en/research/computational-molecular-design/software-server/flexx-molecular-docking.html*.

72.Surflex-Dock. *http://www.tripos.com/index.php?family=modules,SimplePage,,,&page=surflex_dock&s=0*.

73.Knowledge and Discovery from Crystal Structure Data. *http://www.ccdc.cam.ac.uk/Solutions/GoldSuite/Pages/GOLD.aspx*.

74. Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. *Journal of Chemical Information and Modeling* **2009**, *49*, 1455-1474.

75. Li, X.; Li, Y.; Cheng, T.; Liu, Z.; Wang, R. Evaluation of the Performance of Four Molecular Docking Programs on a Diverse Set of Protein-Ligand Complexes. *Journal Computational Chemistry* **2010**, *31*, 2109-2125.

76. GOLD Suite Solutions. *http://www.ccdc.cam.ac.uk/Solutions/GoldSuite/Pages/GOLD.aspx*.

77. Liebeschuetz, J. W.; Cole, J. C.; Corb, O. Pose Prediction and Virtual Screening Performance of GOLD Scoring Functions and a Standardized Test. *Journal of Computer-Aided Molecular Design* **2012**, *26*, 737-748.

78. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical Scoring Functions: The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes. *Journal of Computer-Aided Molecular Design* **1997**, *11*, 425-445.

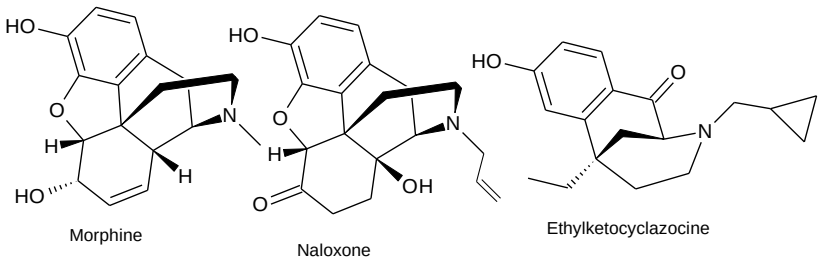
79.Control of the Hydrogen Ion Activity (pH) in the Body. *http://dwb4.unl.edu/Chem/CHEM869R/CHEM869RLinks/www.usyd.edu.au/su/anaes/lectures/acidbase_mjb/control.html;*.

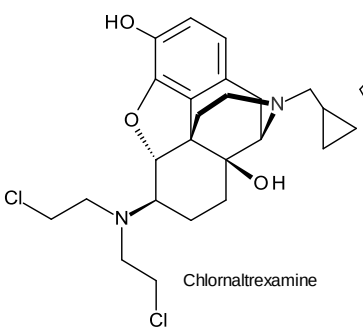
80. Chen, C.; Xue, J.; Zhu, J.; Chen, Y.; Kunapuli, S.; de Riel, J. K.; Yu, L.; Liu-Chen, L. Characterization of Irreversible Binding of β-Funaltrexamine to the Cloned Rat μ Opioid Receptor. *Journal of Biological Chemistry* **1995**, *270*, 17866-17870.

81. Borics, A.; Toth, G. Structural Comparison of μ-Opioid Receptor Selective Peptides Confirmed Four Parameters of Bioactivity. *Journal of Molecular Graphics and Modelling* **2010**, *28*, 495-505.

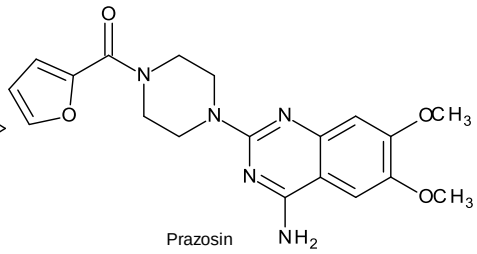
82. Liebmann, C.; Schnitter, M.; Hartrodt, B.; Born, I.; Neubert, K. Structure-Activity Studies of Novel Casomorphin Analogues: Binding Profiles Towards m1-, m2-, and d-Opioid Receptors. *Die Pharmazie* **1991**, *46*, 345-348.

## Appendix

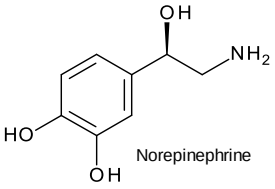




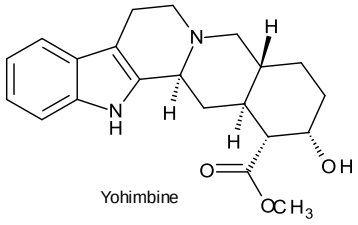

Chlornaltrexamine



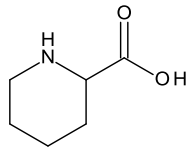

Prazosin




Norepinephrine




Yohimbine




Pipecolic Acid